UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS

DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR

ESTUDIOS PRECLÍNICOS PARA EL TRATAMIENTO MEDIANTE TERAPIA GÉNICA DE LA DEFICIENCIA DE PIRUVATO QUINASA

PRECLINICAL STUDIES FOR THE GENE THERAPY TREATMENT OF PYRUVATE KINASE DEFICIENCY

TESIS DOCTORAL DE:

MARÍA GARCÍA GÓMEZ

BAJO LA DIRECCIÓN DE:

JOSÉ CARLOS SEGOVIA SANZ

SUSANA NAVARRO ORDÓÑEZ

Madrid, 2013

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR

PRECLINICAL STUDIES FOR THE GENE THERAPY TREATMENT OF PYRUVATE KINASE DEFICIENCY

MARÍA GARCÍA GÓMEZ

TESIS DOCTORAL
MADRID 2013

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR

ESTUDIOS PRECLÍNICOS PARA EL TRATAMIENTO MEDIANTE TERAPIA GÉNICA DE LA DEFICIENCIA DE PIRUVATO QUINASA

MARÍA GARCÍA GÓMEZ

TESIS DOCTORAL
MADRID 2013

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR

ESTUDIOS PRECLÍNICOS PARA EL TRATAMIENTO MEDIANTE TERAPIA GÉNICA DE LA DEFICIENCIA DE PIRUVATO QUINASA

Memoria presentada por **MARÍA GARCÍA GÓMEZ** para optar al grado de Doctor Europeo por la Universidad Complutense de Madrid.

Directores de Tesis:

José Carlos Segovia Sanz

Susana Navarro Ordóñez

MARÍA GARCÍA GÓMEZ

TESIS DOCTORAL
MADRID 2013

Dr. José Carlos Segovia Sanz, Jefe de la Unidad de Diferenciación y Citometría de la División de Terapias Innovadoras en el Sistema Hematopoyético del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) y **Dra. Susana Navarro Ordóñez**, investigadora postdoctoral en la División de Terapias Innovadoras en el Sistema Hematopoyético del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), certifican que la memoria adjunta, titulada "ESTUDIOS PRECLÍNICOS PARA EL TRATAMIENTO MEDIANTE TERAPIA GÉNICA DE LA DEFICIENCIA DE PIRUVATO QUINASA" ha sido realizada por la licenciada Dña. María García Gómez bajo la dirección de los que suscriben, y cumple con las condiciones exigidas para optar al título de Doctor Europeo por la Universidad Complutense de Madrid.

José Carlos Segovia Sanz

Susana Navarro Ordóñez

El presente trabajo de investigación ha sido realizado en la División de Terapias Innovadoras en el Sistema Hematopoyético del **CIEMAT** y Centro de Investigaciones Biomédicas en Red de Enfermedades Raras (**CIBERER**) del Instituto de Salud Carlos III, con la colaboración del Proyecto **PERSIST** del programa de salud de la Unión Europea, la Red Española de Terapia Celular (**TERCEL**) y el Ministerio de Economía y Competitividad (**MICINN**), Secretaría de Estado de Investigación, Desarrollo e Innovación.

María García Gómez ha disfrutado de dos contratos como licenciada concedidos por el Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (**CIEMAT**) y el Centro de Investigaciones Biomédicas en Red de Enfermedades Raras (**CIBERER**).

A mis padres, Marisa y Alfonso

A Matt

Acknowledgements

I would like to thank my supervisors, José Carlos Segovia and Susana Navarro, for supporting me through my work and for their guidance throughout these last six years.

Thanks also to Juan Bueren for allowing me to be a part of the Hematopoiesis group during such an exciting time and for encouraging every step of my work.

I am grateful to Michael Antoniou for welcoming me into his group and for his help with the development of erythroid-specific vectors. Also, Axel Schambach for kindly designing the codon optimized RPK transgene.

Thanks to María García-Bravo for her expert advice about cloning experiments, Miguel Ángel Martín for his work with the animals, Israel Orman for his technical assistance, Rebeca Sánchez for helping me to tackle the flow cytometer and Óscar Quintana for all his ideas and help with histology.

I am especially thankful to Matthew Wiseman for carefully checking all the English from this manuscript and Blanca Gómez for the cover design.

Last but not least, thanks very much to all my colleagues and friends from the Innovative Therapies in the Hematopoietic System Division at CIEMAT, for their support in the bad times and their enthusiasm in the good. Thanks for your help, advice and enjoyable times. Thanks also to our neighbours in the Epithelial Biomedicine Division, for your interest and good moments.

Finally I would like to thank my family and friends for their unconditional support. Without doubt, this work would not have been possible without you.

INDEX

ABBREVIATIONS

2,3-DPG	2,3-diphosphoglycerate
3-PG	3-phosphoglycerate
A	Adenine
AAV2	Adeno-associated virus 2
ADA	Adenosine deaminase
ADP	Adenosine diphosphate
ALD	Adenoleukodystrophy
ATP	Adenosine triphosphate
BFU-E	Burst-forming unit-erythroid
BM	Bone marrow
BMP4	Bone morphogenetic protein 4
BMT	Bone marrow transplantation
BSA	Bovine serum albumin
C	Cytosine
CA	Capsid protein
CEP	Congenital erythropoietic porphyria
CFU-E	Colony-forming unit-erythroid
CFU-G	Colony-forming unit-granulocyte
CFU-GM	Colony-forming unit of granulocyte-macrophage
CFU-GMM	Colony -forming unit of granulocyte, macrophage, erythroid and megakaryocyte
CFU-M	Colony-forming unit-macrophage
CFUs	Colony-forming units
CGD	Chronic granulomatous disease
cHS4	Chicken hypersensitive site 4
CIDs	Chemical inducers of dimerization
CIS	Common integration site
CMV	Cytomegalovirus
CNSHA	Chronic non-spherocytic haemolytic anemia
coRPK	Optimized version of the *PKLR* cDNA
cPPT	Central polypurine tract
CTS	Central terminal sequence
DAPI	4',6-diamidino-2-phenylindole
DLA	Dog leukocyte antigen
DLI	Donor lymphocyte infusion
Dpt	days post-transplantation
dsDNA	Double-stranded DNA
EDTA	Ethylendiamintetracetic acid
EF1α	Elongation factor 1α
EGFP	Enhanced Green Fluorescent Protein
Epo	Erythropoietin
EpoR	Erythropoietin receptor
EPP	Erythroid protoporphyria
FBP	Fructose-1,6-diphosphate
FBS	Foetal bovine serum
FECH	Ferrochelatase
FITC	Fluorescein isothiocyanate
FOXO3	Forkhead box O3
G	Guanine
G6PD	Glucose-6phosphate dehydrogenase
GALV	Gibbon ape leukaemia virus
GPI	Glucose-phosphate isomerase
GPx	Glutathione peroxidase
GSH	Reduced glutathione
GSR	Glutathione reductase
GSSG	Oxidized glutathione
GvHD	Graft versus host disease
h*ALAS*	human 5-aminolevulinate synthase gene
hAlb	Human albumin
HBS	HEPES-Buffered saline
HBSS	Hanks' Balanced Salt Solution
HCV	Mean corpuscular hemoglobin
HEK	Human embryonic kidney

HFE	Hereditary hemochromatosis
HFV	Foamy vectors
HGB	Hemoglobin
HIV-1	Human immunodeficiency virus type 1
HK	Hexokinase
HLA	Human leukocyte antigen
HRP	Horseradish peroxidase
HS	DNase I hypersensitive site
HSC	Hematopoietic stem cell
HTC	Hematocrit
IGF-1	Insulin-like growth factor
IL3	Interleukine 3
IL6	Interleukine 6
IMDM	Iscove modified Dulbecco medium
IN	Integrase
IP	Propidium iodide
IRES	Internal ribosome entry site
LCR	Locus control region
LCMS	Liquid chromatography mass spectrophotometry
LDH	Lactate dehydrogenase
LPK	Liver pyruvate kinase
LPK	Liver pyruvate kinase isoform
LSK	Lin^- $Sca1^+$ $cKit^+$ population
LTR	Long terminal repeats
LV	Lentivirus
M1PK	Muscular pyruvate kinase isoform
M2PK	Fetal pyruvate kinase isoform
MA	Matrix protein
MCH	Mean corpuscular hemoglobin
MCV	Mean corpuscular volume
MEL	Mouse erythroleukemia cells
MGMT	Methylguanine methyltransferase
MLD	Metachromatic leukodystrophy
MLV	Moloney mouse leukaemia virus
MOI	Multiplicity of infection
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide reduced
NADP	Nicotinamide adenine dinucleotide-phosphate
NADPH	Nicotinamide adenine dinucleotide-phosphate reduced
NC	Nucleocapsid protein
pA	polyadenilation signal
PBS	Primer binding site
PCR	Polymerase chain reaction
PE	Phycoerythrin
PEG	Pegylated
PEP	Phosphoenolpyruvate
PFK	Phospho-fructokinase
PGK	phosphoglyceratokinase
PGM	Phosphoglycerate mutase
Phe	Phenylalanine
PK	Pyruvate kinase
PKD	Pyruvate kinase deficiency
Pklr	Red cell and liver pyruvate kinase gene
PKM	Muscular pyruvate kinase gene
PR	Protease
qPCR	quantitative polymerase chain reaction
RBC	Red blood cells
RCL	Replication-competent lentiviruses
RE	Regulatory element
rEpo	Recombinant erythropoietin
RFP	Red fluorescent protein

rh-G-CSF	Recombinant human granulocyte colony stimulating factor
rhIL-6	Recombinant human interleukin 6
rmIL-3	Recombinant murine interleukin 3
rmSCF	Recombinant murine stem cell factor
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPK	Red cell pyruvate kinase
RRE	Rev responsive element
RSV	Rous sarcoma virus
RT	Reverse transcriptase
S/MARs	Scaffold/matrix attachment regions
SA	Splice acceptor
SAV	Streptavidin
SCD	Sickle cell disease
SCF	Stem cell factor
SCID	Severe combine immunodeficiency
SD	Splice donor
SEM	Stand error of the media
SFFV	Spleen focus-forming virus
SIN	Self-inactivating
SIV	Simian immunodeficiency virus
SOD	Superoxide dismutase
ssDNA	Single-stranded DNA
SU	Surface protein
SV40	Simian virus 40
T	Thymine
TE	Tris-EDTA buffer
Tfr1	Transferrin receptor
Tm	Melting temperature
TM	Transmembrane protein
TMB	Tetramethylbenzidine
TPI	Triosephosphate isemerase
TRC	Tricolor
TTN	Titin gene
UCOE	Ubiquitous chromatin opening elements
UROS	Uroporphyrinogen III synthase
USE	Upstream sequence element
UTR	Untranslated region
VCN	Vector copy number
VITA	Vector Integration Tag analysis
VSV-G	Vesicular stomatitis virus glycoprotein
WAS	Wiskott-Aldrich syndrome
WBC	White blood cells
WHV	woodchuck hepatitis virus
Wpre	Woodchuck hepatitis virus post-transcriptional regulatory element
X-CGD	X-linked Chronic granulomatous disease
β-LCR	Locus control region from β-Globin gene
γ-RV	Gammaretrovirus
γ-RV	γ-retrovirus
Ψ	Packaging signal

I. RESUMEN

INTRODUCCIÓN:

La deficiencia en Piruvato Quinasa eritrocitaria (PKD) es una enfermedad autosómica recesiva producida por mutaciones en el gene *PKLR* que afecta al metabolismo energético del eritrocito. La Piruvato Quinasa (PK) cataliza la última reacción de la ruta glicolítica, produciendo piruvato y ATP como productos finales a partir de fosfoenolpiruvato (PEP) y ADP. Por tanto es una enzima clave en el metabolismo de carbohidratos (Figura 1). La PK se expresa en todas las células vivas del organismo, y debido a su importante papel en el metabolismo celular está regulada por un estricto control alostérico en respuesta a distintos efectores (ej. ATP actúa como inhibidor y fructosa-1,6-difosfato como activador). Además, la expresión de PK está sujeta a un control temporal y espacial por mecanismos específicos, para adaptar su actividad enzimática a las necesidades metabólicas particulares de cada tejido y etapa de diferenciación o desarrollo. En humanos se han identificado cuatro isoenzimas específicas de tejido: PKM1 (tejidos adultos, ej. músculo esquelético, corazón y cerebro), PKM2 (tejidos proliferativos, tanto fetales como tumorales), LPK (hígado, córtex renal, intestino delgado y células β pancreáticas) y RPK (células eritroides), codificadas por los genes estructurales, PKM (cromosoma 15, isoformas M1 y M2) y *PKLR* (cromosoma 1, isoformas L y R).

La secuencia codificante del gen *PKLR* está dividida en 12 exones, 10 de cuales son compartidos por las dos isoformas (Figura 4). Los dos primeros exones son específicos de cada isoenzima: 1(R) para la isoforma eritroide y 2(L) para la isoforma hepática, también llamado exón 1(L). A pesar de que las mutaciones en el gen *PKLR* podrían afectar también a los hepatocitos, los síntomas clínicos de la PKD están restringidos a las células rojas de la sangre (RBCs), ya que la deficiencia hepática es normalmente compensada por una síntesis persistente de LPK en los hepatocitos. La PK muestra también una patrón de expresión diferencial a lo largo del proceso de maduración de las células eritroides. Así, durante la eritropoyesis la expresión de PKM2 decae gradualmente a partir del estadio de eritroblasto basófilo, y la síntesis de RPK aumenta simultáneamente a través de la activación de su promotor específico (promotor PKLR, Figura 5) por factores de transcripción específicos de tejido eritroide.

Los eritrocitos utilizan principalmente la glicolisis, la ruta de las pentosas fosfato y el ciclo del glutatión para cubrir sus necesidades metabólicas. Sin embargo, debido a su alto grado de especialización, son incapaces de sintetizar proteínas *de novo* o de metabolizar la

glucosa vía fosforilación oxidativa en las mitocondrias, ya que las RBCs maduras carecen de núcleo, mitocondrias y retículo endoplasmático (Figura 3). De esta manera, los eritrocitos dependen de la glicolisis casi exclusivamente para la producción de ATP y para el mantenimiento de su integridad celular. La presencia de RPKs no funcionales en las células eritroides conlleva dos problemas metabólicos fundamentales: la reducción de los niveles de ATP y el aumento de 2,3-difosfoglicerato (2,3-DPG), alterando la homeostasis y reduciendo la vida media de los eritrocitos deficientes. Aunque el mecanismo exacto por el cual la deficiencia de RPK lleva a una reducción de la supervivencia de las RBCs en circulación se desconoce, tanto alteraciones en las propiedades de membrana debidas a un desequilibrio en los niveles intracelulares de K^+ y $Ca2^+$, como el bloqueo metabólico causado por la acumulación de metabolitos intermediarios de la glicolisis han sido propuestos. Este desequilibrio iónico y metabólico en las RBCs deficientes conlleva alteraciones graves en su funcionalidad, causando hemolisis intravascular y destrucción prematura de los eritrocitos deficientes en el bazo.

La deficiencia de RPK es el defecto enzimático hereditario más frecuente de la ruta glicolítica en los eritrocitos, causando anemia hemolítica crónica no esferocítica (CNSHA). Hasta el momento se han descrito más de 400 casos graves, identificando la presencia de un gran número y variedad de mutaciones en el gen *PKLR*. Las aproximadamente 195 mutaciones diferentes asociadas con CNSHA están repartidas por toda la región codificante de *PKLR*, y la mayoría son mutaciones por cambio de sentido (69%), mutaciones de *splicing* (13%) y mutaciones sin sentido (5%), aunque pequeñas deleciones, inserciones y cambios en el marco de lectura han sido también identificadas (Tabla 1). La deficiencia de RPK se transmite de manera autosómica recesiva, presentando una amplia distribución tanto geográfica como étnica. Aunque la prevalencia de la PKD todavía no se conoce con exactitud, se ha estimado en 1:20.000 casos en la población caucásica de Norte América.

Las manifestaciones patológicas de la PKD normalmente aparecen cuando la actividad enzimática es menor del 25% de la actividad RPK normal, afectando a pacientes homocigotos o heterocigotos compuestos en los que cada uno de los dos alelos porta una mutación diferente. Debido a la gran heterogeneidad genética y molecular de la patología PKD, la sintomatología varía notablemente desde una anemia hemolítica leve o compensada, a una anemia grave que requiere transfusiones sanguíneas desde el nacimiento y durante toda la vida del paciente. Además, en algunos casos se ha demostrado

la a falta de expresión compensatoria de la isoenzima PKM2 en eritrocitos, causando un fenotipo grave. La PKD también cursa con elevada reticulocitosis y esplenomegalia en los casos más graves, y tanto retraso en el crecimiento como muerte *in utero* o durante el periodo neonatal han sido descritos en casos poco frecuentes. No obstante, los síntomas hematológicos observados en los pacientes no son suficientemente específicos de esta patología, y ni la actividad enzimática residual en RBCs ni el número de reticulocitos en circulación están relacionados con la gravedad del cuadro hemolítico, complicando el diagnóstico de la deficiencia de PK.

Hasta el momento no existe una terapia efectiva disponible para los pacientes más graves de PKD, y las medidas paliativas son la única opción. Normalmente los paciente reciben transfusiones sanguíneas periódicas que a largo plazo disminuyen considerablemente su calidad de vida, causando importantes problemas por la sobrecarga de hierro en el hígado y haciendo necesario un tratamiento adicional con agentes quelantes de hierro. La esplenectomía ha reportado un beneficio clínico en algunos pacientes graves de PKD, ya que aunque no revierte la hemolisis, resulta en un aumento de los niveles de hemoglobina de 1-2 g/dL que suprime la necesidad de transfusiones en la mayoría de los pacientes. Sin embargo, los pacientes esplenectomizados son más susceptibles a padecer infecciones y trombos, existiendo también riesgo de aplasia y crisis hemolíticas. El trasplante alogénico de médula ósea (BM) ha sido utilizado en ciertos pacientes graves, siendo exitoso en uno de ellos. Sin embargo, la disponibilidad de donantes HLA-idénticos limita el número de pacientes que pueden ser sometidos a trasplante de BM, y tanto los tratamiento previos de acondicionamiento como la inmunosupresión posterior pueden causar problemas de toxicidad incluso más graves que la propia deficiencia enzimática.

La terapia génica basada en el trasplante de células madre hematopoyéticas autólogas genéticamente corregidas evita estos inconvenientes, siendo teóricamente aplicable a todos los pacientes. Además, no existe riesgo de enfermedad de injerto contra huésped (GvHD) ni necesidad de inmunosupresión post-trasplante. Este procedimiento terapéutico consiste en la introducción de una versión salvaje del gene responsable de la enfermedad en las células diana apropiadas, las cuales repoblarán todos los linajes hematopoyéticos en el paciente suministrando células sanguíneas funcionales capaces de revertir el fenotipo de la enfermedad. En el caso de enfermedades monogénicas como la PKD, el éxito de la terapia génica requiere la expresión estable y a largo plazo del transgén

en las células afectadas, siendo los vectores integrativos los vehículos más adecuados para conseguir una expresión eficiente. De entre ellos, los vectores γ-retrovirales (γ-RV) han sido ampliamente utilizados con éxito en modelos preclínicos. Sin embargo, la ocurrencia de leucemias T en los primeros ensayos clínicos con pacientes de inmunodeficiencia combinada severa ligada al cromosoma X (X1-SCID), causadas por la integración del vector γ-retroviral en el genoma de la célula diana produciendo la desregulación de la expresión de oncogenes (mutagénesis insercional), resaltó la necesidad de desarrollar vectores más eficientes y seguros. Los resultados procedentes de numerosos ensayos preclínicos (ej. β-talasemia, X1-SCID, Wiskott-Aldrich syndrome -WAS-, leucodistrofia metacromática -MLD-, anemia de Fanconi) apoyan el uso de vectores lentivirales autoinactivantes (LV), que en la actualidad ofrecen el mejor equilibrio entre eficacia y seguridad de la terapia génica. De hecho, actualmente existen varios protocolos clínico en marcha para el tratamiento de pacientes de WAS, Anemia de Fanconi y adenoleucodistrofia (ALD).

Tanto los γ-retrovirus como los lentivirus pertenecen a la familia *Retroviridae*, y presentan unas características particulares que los hacen especialmente adecuados como vectores de terapia génica: 1) su ciclo viral de ambos requiere la transcripción reversa e integración del genoma viral en el genoma de la célula huésped, proporcionando una expresión estable y a largo plazo de los genes virales en las células infectadas y su progenie (Figura 6). 2) Su genoma ha sido ampliamente caracterizado, haciendo posible su modificación. 3) Las proteínas virales estructurales pueden ser aportadas en *trans*, asegurando la expresión ectópica sólo de los transgenes, y aumentando la seguridad de los protocolos de terapia génica, ya que las proteínas virales no se expresan después de la transducción. 4) Las proteínas de la envuelta viral pueden ser también modificadas, haciendo posible el desarrollo de envueltas quiméricas para redirigir el tropismo del virus. 5) Los LV, a diferencia de los γ-retrovirus, son capaces de infectar tanto células en división como células quiescentes. Además, mientras que los γ-retrovirus tienen una preferencia de integración por las secuencias promotoras de genes activos, 6) los LV se integran preferentemente en las unidades transcripcionales activas, presentando un menor potencial oncogénico.

La mayoría de vectores LV validados en protocolos preclínicos están basados en el virus HIV-1, cuyo genoma se organiza en tres genes estructurales: *Gag, Pol* y *Env* (Figura 7). Además de las proteínas codificadas por estos genes y compartidas con los γ-RV, el genoma

de los LV también incluye otras proteínas accesorias (*Tat, Rev, Nef,* etc.) siendo su genoma más complejo. Los LV son capaces de portar transgenes relativamente grandes (hasta 10 kb). En general, la generación de partículas LV se consigue por la cotransfección de tres plásmidos en la línea humana productora transitoria 293T: el plásmidos empaquetador, el plásmido de transferencia y el plásmido de envuelta (Figura 7). Dado que el HIV-1 es un patógeno humano, los métodos de producción de vectores LV están dirigidos a reducir las secuencias homólogas entre el vector de transferencia y las secuencias accesorias (genes estructurales y reguladores), separando estas funciones en distintos plásmidos, evitando así la generación de LV competentes en replicación. Además, se han introducido numerosas modificaciones, tanto en el plásmido de transferencias como en las secuencias accesorias, con el fin de aumentar su seguridad y optimizar la expresión del transgén.

En cuanto a seguridad destacan las modificaciones de las regiones LTR. Así, los vectores LV autoinactivantes (SIN) poseen una deleción de 400 pb en la región U3 de la secuencia 3'LTR que cubre las secuencias promotoras/enhancer del virus. De manera que la transcripción del transgén está dirigida por el promotor interno incluido en el plásmido de transferencia (Figura 8). Esto hace posible la elección de promotores internos débiles o con actividad moderada para conseguir niveles de expresión del transgén adecuados y/o fisiológicos, y al mismo tiempo disminuir la probabilidad de transactivación de genes vecinos. Por otra parte, la región promotora U3 de la secuencia 5'LTR se ha sustituido por promotores heterólogos (ej. CMV y RSV) para aumentar el título de los sobrenadantes virales producidos. Esta secuencia dirige la expresión de los transcritos primarios en la célula empaquetadora, pero no se integra en el genoma de las células transducidas (Figura 7), ya que después de la transcripción reversa e integración del provirus, la unidad transcripcional queda flanqueada por la secuencia 3'LTR en ambos extremos.

Como se ha mencionado anteriormente, la genotoxicidad derivada del uso de vectores integrativos es uno de los principales inconvenientes en el campo de la terapia génica, debido a los potenciales eventos leuquemogénicos debidos de la integración del vector en el genoma de los progenitores hematopoyéticos. Este fenómeno se conoce como mutagénesis insercional y puede causar la desregulación de genes próximos (Figura 9). Las mutaciones más peligrosas son las que implican la activación de protooncogenes próximos al sitio de integración o de genes del ciclo celular, así como la inhibición de genes supresores de tumores. Diversos estudios han demostrado que los LV tienen una preferencia de

integración por las unidades transcripcionales activas, mientras que los γ-RV se integran preferentemente en regiones reguladoras de genes activos o en lugares adyacentes de las secuencias hipersensibles a DNasaI. De manera que los LV presentan un menor potencial oncogénico, y por ello en la actualidad son las herramientas de elección en los protocolos de terapia génica dirigidos a corregir enfermedades genéticas monogénicas.

A pesar de que los LV presentan un menor potencial de mutagénesis insercional, las secuencias polyA contenidas en su región 3'LTR todavía podrían causar la transcripción aberrante del transgén por lectura transcripcional continuada (*readthrough*), y el promotor interno podría transactivar promotores celulares próximos por un fenómeno de interferencia del promotor. En este sentido, el uso de promotores específicos de tejido ofrece beneficios adicionales a la seguridad de la terapia génica, disminuyendo los posibles efectos adversos en células no afectadas. La mayoría de promotores y secuencias potenciadoras de expresión (*enhancers*) específicos de tejidos eritroide han sido validados en modelos de hemoglobinopatías, siendo el promotor y las secuencias reguladoras del gen humano de la β globina (*Locus control regions* -βLCR-) los más utilizados (Tabla 3). Las secuencias βLCR están formadas por elementos reguladores específicos de tejido eritroide que confieren niveles de expresión fisiológicos. Su mecanismos de regulación no se conoce totalmente, pero se piensa que actuarían haciendo la cromatina más accesible para los factores de transcripción, tanto ubicuos como específicos del linaje eritroide, siendo su mecanismo dependiente del número de copias integradas.

La demostración de que la terapia génica podría ser una alternativa terapéutica para la PKD ha sido demostrada en trabajos previos, pero la eficiencia de transducción y el porcentaje de quimerismo donante siguen siendo las mayores limitaciones. Además, la reducción del riesgo de mutagénesis insercional requiere el desarrollo de nuevos vectores, y los beneficios evidentes del uso de promotores internos apuntan a un diseño lentiviral autoinactivante como vector ideal para la terapia génica de PKD, incluyendo un promotor específico de tejido eritroide, o al menos un promotor interno débil en su secuencia para dirigir la expresión de la RPK humana.

JUSTIFICACIÓN Y OBJETIVOS:

La herencia autosómica recesiva que caracteriza la deficiencia de RPK, junto con la restricción del defecto enzimático a las células eritroides, hacen de la PKD una enfermedad adecuada para ensayar protocolos de terapia génica. Además, la falta de opciones terapéuticas satisfactorias para pacientes con lesiones graves en el gen *PKLR* hacen especialmente interesante el desarrollo de nuevas protocolos de terapia génica desde un punto de vista traslacional. Sin embargo, el escaso número de pacientes PKD con fenotipo grave que requieren trasplante de BM, y la falta de una aparente ventaja selectiva de las células genéticamente corregidas frente a las deficientes, dificultan el desarrollo de nuevas estrategias terapéuticas basadas en la corrección génica del gen afectado, siendo el número de estudios preclínicos todavía muy limitado.

Con el objetivo de esclarecer los mecanismos fisiopatológicos implicados en el desarrollo de la deficiencia de RPK, y de desarrollar un tratamiento eficaz basado en la modificación genética de células madre hematopoyéticas deficientes, en este trabajo se consideraron los siguientes objetivos parciales:

1. Estudiar las características fenotípicas del modelo de ratón AcB55 deficiente en *Pklr*, para determinar su utilidad como modelo experimental de la patología PKD en humanos.
2. Desarrollar un vector lentiviral de potencial aplicación clínica con expresión ubicua del transgén humano *PKLR wild type*, para llevar a cabo protocolos preclínicos de terapia génica y evaluar así la eficacia y seguridad del vector desarrollado.
3. Desarrollar vectores lentivirales con expresión específica de tejido eritroide, proporcionando nuevos vectores de expresión restringida para futuros protocolos de terapia génica para PKD.

RESUMEN DE RESULTADOS:

- Los ratones AcB55 reproducen los síntomas hematológicos de la enfermedad PKD humana, mostrando una reducción significativa en el recuento de RBCs, niveles de hemoglobina e índice de hematocrito (Figura 18), junto con esplenomegalia aguda (Figura 20) y reticulocitosis constitutiva (Figura 19). Además, los ratones AcB55

muestran una intensa sobrecarga de hierro en el hígado (Figura 24.B), reproduciendo uno de los principales problemas de los pacientes gravemente afectados por la deficiencia de RPK.

- La mutación con pérdida de función 269 T>A en el gen *Pklr* que portan los ratones AcB55 conlleva la reducción de la vida media de las RBCs (Figura 21), desencadenando mecanismos compensatorios (eritropoyesis extramedular o estresada) (Figuras 23 y 24.A) que implican el aumento los niveles de Epo (controles Figuras 31 y 45) para aminorar la anemia hemolítica. La deficiencia de RPK no sólo afecta a la supervivencia de las RBCs, sino también a la maduración de progenitores eritroides en la médula ósea y el bazo (Figura 22), lo que resulta en una eritropoyesis ineficiente que impide el aumento en la producción de eritrocitos maduros.

- Los ratones PKD mostraron una reducción en el número de progenitores hematopoyéticos pluripotentes (LSK y Slam) tanto en la médula ósea como en el bazo (Figura 26.B), junto con una reducción de progenitores CFUs en médula ósea (Figura 25.A) sin alterar la celularidad de este órgano (Figura 26.C). Sin embargo, tanto la celularidad como el contenido de progenitores CFUs aumenta en el bazo (Figura 25.A y 26.C), demostrando una eritropoyesis extramedular activa. Además, la deficiencia de RPK conlleva una respuesta de migración de progenitores desde la médula ósea a la periferia alterada tras la estimulación con G-CSF (Figura 25.B).

- La aproximación de terapia génica basada en el vector lentiviral pRRL-CMV-hRPK desarrollado proporcionó una corrección parcial del fenotipo en ratones PKD trasplantados (Figuras 28-33) debido al bajo porcentaje de transducción y al silenciamiento del promotor CMV (Tabla 8).

- El protocolo de terapia génica basado en el vector lentiviral pCCL-hPGK-coRPK desarrollado revirtió eficientemente la patología PKD en ratones deficientes trasplantados, tanto primarios (Figuras 35 y tabla 10) como secundarios (Figuras 46, 48-52) con una dosis génica de 1.65 copias del provirus integrado por célula (Tabla 11). El promotor humano PGK fue suficientemente fuerte como para expresar niveles terapéuticos de la proteína humana RPK cuando el nivel de quimerismo

donante fue ≥ 60% (Tabla 9), normalizando los parámetros hematológicos (Figura 35 y tabla 10) y los niveles de reticulocitos en periferia (Figura 36) , extendiendo la vida media de los eritrocitos (Figuras 37 y 38) y revirtiendo el mecanismo de eritropoyesis compensatoria (Figuras 39-43).

- La reversión de la anemia hemolítica como consecuencia de la corrección genética con el vector pCCL-hPGK-coRPK rindió niveles normales de Epo (Figura 45), disminuyendo el número de progenitores CFUs en el bazo hasta valores control (Figura 44) y rescatando la patología en los órganos, incluyendo la reversión de la eritropoyesis extramedular y la sobrecarga de hierro (Figuras 40 y 41).

- La expresión ectópica de la proteína RPK humana corrigió la deficiencia energética en los eritrocitos de ratones deficientes (Figuras 53-55) sin alterar el equilibrio metabólico de los leucocitos (Figura 56), subrayando la eficacia y seguridad de este protocolo de terapia génica.

- Los vectores lentivirales que portan el promotor de la β globina humana o el promotor PKLR, fueron altamente específicos de tejido eritroide en experimentos *in vitro* de transducción (Figuras 57 y 58). La combinación de elementos HS3-HS2 con el promotor de la β globina humana fue la que proporcionó una expresión más elevada del transgén coRPK (Figura 59), pero el promotor PKLR mostró una elevada especificidad eritroide, incluso en ausencia de las secuencias potenciadoras de βLCR (Figura 58). Además, la expresión ectópica de coRPK fue superior en células eritroide HEL que en la línea celular eritroide K562 (Figura 59), sugiriendo una expresión más fisiológica del transgén coRPK cuando fue dirigida por su propio promotor.

CONCLUSIONES:

1. Los ratones AcB55, portadores de la mutación con pérdida de función 269 T>A en el gen *Pklr*, constituyen un modelo adecuado de la deficiencia de piruvato quinasa humana, reproduciendo los principales síntomas observados en pacientes de PKD, permitiendo el estudio de los mecanismos fisiopatológicos implicados en la patología PKD y proporcionando un escenario ideal para el desarrollo de estrategias de terapia génica.

2. El vector lentiviral pCCL-hPGK-coRPK desarrollado que porta el promotor humano ubicuo PGK y la secuencia de codones optimizada correspondiente al cDNA del gen humano *PKLR*, revierte eficazmente el fenotipo anémico de los ratones AcB55 mediante transducción y trasplante de progenitores hematopoyético deficientes, siendo, por tanto, un vector de terapia génica eficiente y seguro para su uso en futuros ensayos clínico dirigidos a pacientes graves de PKD que no disponen de un donante HLA-compatible.

3. Hemos construido nuevos vectores lentivirales con promotores específicos de tejido eritroide: el promotor del gen de la β-globina humana y el promotor PKLR, que fueron altamente específicos en estudios *in vitro* de transducción. El promotor PKLR dirigió la expresión del transgén coRPK de una manera específica del tejido eritroide, incluso en ausencia de las secuencias *enhancers* de βLCR, proporcionando un nuevo vector fisiológicamente regulado que podría resultar de utilidad en futuras aproximaciones de terapia génica para la PKD.

APORTACIONES FUNDAMENTALES:

1. A diferencia de otras enfermedades como la talasemia mayor o la anemia falciforme, la deficiencia de RPK no dispone de modelos animales transgénicos, aunque existen varios modelos mutantes espontáneos de *Pklr,* tanto de perro como de ratón. De entre ellos, los ratones AcB55 fueron los últimos en ser encontrados (Min-Oo et al., 2003), por lo que no se disponía de una información completa sobre su fenotipo PKD. En este trabajo, hemos estudiado en detalle el fenotipo hematológico de los ratones AcB55, abarcando nuevas características de la patología PKD no estudiadas hasta la fecha. Nuestros resultados demuestran la utilidad de los ratones AcB55 como modelo experimental de la PKD, ideal para el ensayo de nuevas aproximaciones terapéuticas, y aportan nuevos datos acerca de los mecanismos fisiopatológicos implicados en esta enfermedad.

2. Las características genéticas de la deficiencia de PKD abren la posibilidad de ser tratada mediante terapia génica. Sin embargo, la falta de ventaja selectiva de las células corregidas sobre las deficientes hace necesario un elevado porcentaje de quimerismo donante para conseguir un beneficio clínico de la terapia génica para

PKD. Además, es importante la elección de un promotor suficientemente fuerte para expresar niveles terapéuticos del transgén en las células rojas maduras, que son las células afectadas por la deficiencia de RPK; pero a la vez capaz de mantener su expresión en un rango fisiológico para evitar efectos adversos, tanto genotóxicos como citotóxicos. El protocolo preclínico de terapia génica basado en el vector lentiviral pCCL-hPGK-coRPK desarrollado durante la realización de este trabajo cumple estos requisitos, proporcionando una corrección eficaz y segura de la patología PKD desde un punto de vista tanto fenotípico como metabólico. Este trabajo proporciona por primera vez un vector lentiviral para el tratamiento de la PKD de potencial uso clínico, que podría ser utilizado en futuros ensayos clínicos una vez validado en progenitores hematopoyéticos de pacientes PKD.

3. La corrección genética de progenitores hematopoyéticos con el vector lentiviral pCCL-hPGK-coRPK revertió la patología en ratones PKD de manera eficiente, segura y a largo plazo, corrigiendo el defecto energético en RBCs sin alterar el equilibrio metabólico en leucocitos. Además, el uso del promotor débil hPGK proporcionó niveles terapéuticos del transgén con una dosis génica comprendida en los estándares clínicos, sin fenómenos de silenciamiento observados. La corrección del fenotipo PKD se mantuvo en receptores secundarios trasplantados, sin presentar efectos adversos asociados a la integración proviral, resaltando la seguridad del protocolo preclínico propuesto. Además, el diseño del vector lentiviral no incluyó genes marcadores en su secuencia ni métodos de selección *in vivo* de las células corregidas, evitando posibles problemas inmunológicos, y por tanto, demostrando el potencial terapéutico de este vector.

4. Este trabajo además aporta nuevos vectores lentivirales altamente específicos de tejido eritroide para PKD, cuyo uso podría evitar potenciales efectos adversos de la terapia génica en otros linajes hematopoyéticos no afectados por la deficiencia. Entre ellos, el vector fisiológicamente regulado pCCL-PKLR-coRPK podría tener ventajas adicionales para su potencial aplicación en pacientes PKD, ya que el transgén está sujeto a una regulación natural por su propio promotor. El estudio y validación de los vectores desarrollados en células primarias de pacientes PKD determinará cuál de los diseños lentivirales propuestos proporcionaría el mejor equilibrio entre eficacia y seguridad para futuros ensayos clínicos.

II. SUMMARY

INTRODUCTION:

Red cell Pyruvate Kinase (RPK) deficiency (PKD) is an autosomal recessive disorder caused by mutations in the *PKLR* gene that produces chronic nonspherocytic hemolytic anemia (CNSHA), being the most frequent erythroid hereditary enzymatic defect. *PKLR* gene encodes two of the four PK isoforms: R-type or RPK, specifically expressed in erythrocytes, and L-type or LPK, expressed in the liver, using alternative first exons as post-transcriptional regulatory mechanisms. Clinical symptoms are restricted to red blood cells (RBCs) because the hepatic deficiency is usually compensated by the persistent synthesis of LPK enzyme in hepatocytes.

Mature erythrocytes rely almost exclusively on glycolysis for producing ATP and maintaining cell integrity, as they loose the mitochondria, the nucleus and the endoplasmic reticulum during their differentiation process. They are unable to synthetize proteins *de novo* or to metabolize pyruvate via oxidative phosphorilation. Therefore, any loss of RPK activity will impair RBC metabolism and lifespan due to two major problems: ATP depletion and glycolytic intermediates accumulation that eventually may lead to apoptosis. PKD pathological manifestations usually appear when enzyme activity falls below 25% of the normal RPK activity, and vary considerably from mild or fully compensated hemolytic anemia, to life threatening neonatal anemia requiring exchange transfusions and subsequent continuous support. These clinical procedures often impair the patient quality of life, and iron overload becomes a serious problem for PKD transfused patients.

PKD is transmitted as an autosomal recessive trait and clinical symptoms usually occur in homozygous or compound heterozygous patients carrying two different mutant alleles. Around 195 different mutations in the *PKLR* gene are associated with CNSHA, located throughout the whole coding region with no preference for the active site. RPK deficiency has a worldwide distribution, and three of the most common mutations (1529A, 1456T and 1468T) have a strong ethnical and regional background. The prevalence of PKD varies considerably, but it has been estimated at 1:20 000 in the general Caucasian population. Nonetheless, PKD prevalence is thought to be higher due to the low survival rate of severely affected patients, leading to underestimate. It is also expected to be higher in communities with high incidences of consanguinity.

Severe disease has been associated with high degrees of reticulocytosis and splenomegaly. Growth retardation, *hydrops foetalis* and death during the neonatal period have also been reported with low frequency in severe patients. Clinical severity of PKD has been also associated with a compensatory persistence of the M2PK isoform expression in RBCs, but this does not occur in all patients, leading to an aggravated phenotype. From a clinical point of view, PKD is a complex disease because of its high molecular heterogeneity and variable symptomatology. These characteristics hamper PKD accurate diagnosis and the choice of an appropriate treatment.

Currently there is no specific therapy available for severe cases of PKD. Allogeneic bone marrow transplantation has been used in limited severe cases, but the availability of human leukocyte antigen (HLA)-compatible donors limits the number of patients that can benefit from this treatment. Gene therapy via autologous hematopoietic stem cell (HSC) transplantation overcomes these limitations, being potentially applicable to every patient and providing a supply of functional blood cells able to rescue the disease. The recessive inheritance trait of PKD, and its curative treatment by allogeneic hematopoietic progenitor transplantation make PKD a suitable disease to be treated by HSC gene therapy. However, the low number of patients with poor prognosis requiring bone marrow transplantation, and the lack of an apparent selective advantage of the corrected cells over the diseased ones turns the development of gene therapy approaches for PKD into a challenging task.

When long-term expression is needed for the genetic correction of a disease, viral integrative vectors are required. Among them, γ-retroviral (γ-RV) vectors have been traditionally used in clinical trials, but the occurrence of unexpected adverse effect due to insertional mutagenesis in several X1-linked severe combined immunodeficiency (X1-SCID) transplanted patients made self-inactivating lentiviral vectors (SIN-LV) replace the γ-RV in the translational gene therapy field. Unlike γ-RV, LV vectors are able to infect both, dividing and quiescent cells, and they display a lower oncogenic risk. While γ-RV have a preference towards integrating near promoters of active genes and around DNase I hypersensitive site (HS), LV are more likely to integrate into active transcription units. Moreover, LV vectors have additional modifications into their long terminal repeat (LTR) sequences to increase their safety, such as the 400 bp deletion in the 3'LTR that inactivates the vectors and allows the use of internal promoters to regulate transgene expression. Also, the U3 region from 5'LTR has been replaced by other promoting sequences (i.e. cytomegalovirus -CMV- or Rous

Sarcoma Virus -RSV- promoters) in order to increase the viral titer during lentiviral production.

Genotoxicity is one of the major limitations in the gene therapy field, due to potential hematological malignancies derived from the vector integration. Either enhancers/promoters present in the vector or aberrant splicing from the vector transcripts can mediate insertional mutagenesis. Despite in SIN-LV designs all LTR-dependent effects are avoided, the residual risk of oncogene transactivation by internal promoters is still present. The use of promoters with moderate to weak activity, such as the human phosphoglyceratekinase (PGK) promoter, significantly decreases the risk of insertional transformation, as well as the use of tissue-specific promoters that provide additional benefits. The proof-of-principle that gene repair strategy could be a feasible option to treat PKD has been demonstrated in previous works, but transduction efficacy and donor chimerism level are still the major limitations.

OBJECTIVES:

1. To study the phenotypic characteristic of a *Pklr* deficient mouse strain (AcB55) in order to determine its usefulness as an experimental model of the human PKD disease.

2. To develop a clinically applicable lentiviral vector with ubiquitous expression of the human wild type *PKLR* therapeutic transgene to perform preclinical gene therapy assays, evaluating the efficacy and safety of this vector.

3. To develop erythroid tissue-specific lentiviral vectors to provide new expression-restricted vectors for future gene therapy protocols for PKD.

RESULTS and CONCLUSIONS:

We first characterized the hematological phenotype of the AcB55 mouse strain in order to have a suitable experimental model to assay further gene therapy protocols. We demonstrated that these mice reproduced the main symptoms of the human PKD disease: hemolytic anemia, reticulocytosis and splenomegaly, together with an intense liver iron

overload, one of the major problems of severely affected or repeatedly transfused PKD patients. The 269 T>A loss-of-function mutation in the *Pklr* gene led to shortened RBC half-life, triggering compensatory mechanisms (extramedullar or stress erythropoiesis and reticulocytosis) involving high levels of Epo to ameliorate the hemolytic anemia. RPK deficiency affected not only the survival of RBCs, but also the maturation of erythroid progenitors in bone marrow and spleen, which resulted in ineffective erythropoiesis failing to increase the production of mature RBCs.

Our study also reveals new insights about PKD pathology, such as an imbalance in the total hematopoietic colony-forming units (CFU) progenitors and uncommitted LSK and Slam progenitors, both in bone marrow and spleen. In addition, PKD mice showed an impaired migration of progenitors from the BM to the periphery upon granulocyte colony stimulating factor (G-CSF) mobilization, which may suggest unexpected additional functions beyond the committed erythroid lineage for the RPK enzyme. The characterization of the AcB55 mouse model allowed the study of the pathological mechanisms involved in PKD pathology, providing an ideal scenario for developing gene therapy strategies.

To address the development of an efficient gene therapy protocol for PKD, we based our strategy on self-inactivating lentiviral vectors (LV) as they offer the best balance between efficacy and safety for HSC genetic modification. Our results demonstrate for the first time a long-term phenotypic correction of PKD disease through LV-based gene therapy. We tested two different ubiquitous promoters, CMV and hPGK, driving the expression of the human *PKLR* cDNA. While CMV promoter was associated with transgene silencing events, human PGK promoter was potent enough to express clinically relevant levels of RPK transgene in both, primary and secondary deficient mice transplanted with transduced cells. With a viral dose of 1.65 copies of the integrated provirus per cell and a donor chimerism ≥ 60%, the genetic correction protocol was able to extend RBC half-life, with the consequent correction of hematological variables and reticulocyte levels, and to revert the compensatory erythropoiesis. In addition, transgenic hRPK expression led to normal Epo levels and spleen CFU progenitor content, with a clear correction of PKD organ pathology. Metabolomic studies corroborated the correction of the energetic defect in RBCs, showing no metabolic disturbances in leukocytes and underscoring the therapeutic potential of the developed hPGK-carrying lentiviral vector.

Additionally, we developed new lentiviral vectors carrying the *PKLR* cDNA optimized transgene (coRPK) and two different erythroid-specific promoters, the β-globin promoter and the regulatory sequences of the *PKLR* gene, to provide new expression-restricted vectors that may be useful for future PKD gene therapy approaches. To foster both specificity and efficacy of the PKLR promoter, we also included the regulatory sequences of the human β-globin locus control region (LCR), achieving high erythroid-specificity for all the vectors developed. Vector carrying the combination of the DNase I hypersensitive sites (HS) HS3 and HS2, together with the β globin promoter provided the highest expression of coRPK transgene in *in vitro* transduction studies, but PKLR promoter showed a high erythroid-specificity even in the absence of βLCR enhancer sequences. In addition, ectopic coRPK expression was higher in HEL erythroleukemic cells than in the erythroid K562 cell line, suggesting a more physiological expression of the coRPK transgene when was driven by its own promoter

Overall, our preclinical results provide encouraging expectations that HSC genetic correction with the hPGK-based lentiviral vector could be a therapeutic option for severely affected PKD patients lacking HLA-compatible donor, paving the way towards the development of future clinical trials. Currently, gene replacement addition strategy appears to be the most reliable and applicable strategy. The gene therapy protocol proposed in this work would be able to correct the wide range of mutations described in PKD patients, including mutations in the regulatory sequences of the gene. On the other hand, as *PKLR* null mutations are rare, most patients display truncated versions of the RPK protein or non-functional complete proteins, avoiding a possible immunological response against the exogenous proteins in patients transplanted with genetically corrected cells.

Nevertheless, the ubiquitous nature of hPGK promoter made us consider a new genetic correction strategy based on enhanced tissue-specific or physiological-regulated vectors. The targeted coRPK expression to the erythroid lineage may lead to a decrease risk of adverse effects in other cell-types not affected by the disease, such as leukocytes or granulocytes. However, as we speculated above, the RPK enzyme may have additional functions beyond the erythroid compartment. Although we are currently exploring this possibility, the demonstration of this idea would support the use of the hPGK ubiquitous promoter in gene therapy vectors for PKD, providing the best balance between efficacy and safety for future clinical trials.

III. INTRODUCTION

1. Pyruvate kinase (PK): structure and function

Pyruvate kinase (PK; EC 2.7.1.40) catalyses the second adenosine triphosphate (ATP)-generating reaction of the glycolysis pathway in the cells by converting the phosphoenolpyruvate (PEP) into pyruvate, requiring Mg^{2+} (or Mn^{2+}) and K^{+} for its activity. PK activity is particularly important in anaerobic glycolysis yielding nearly 50% of the total ATP production (Zanella et al., 2007), and serves as a critical checkpoint for glycolytic intermediate levels to support anabolic reactions in the cells. The substrate (PEP) and the final product (pyruvate) of the PK-catalysed reaction are connected with many other biosynthetic pathways (nucleotides, carbohydrates, amino acids and lipids) (Figure 1), and both strict regulation and optimal activity of the PK enzyme are greatly important for the entire cellular metabolism.

Because of its central role in metabolism, PK is tightly regulated by allosteric responses to different effectors. PK displays a positive cooperativity towards its substrate PEP and is allosterically activated by fructose-1,6-diphosphate (FBP), while ATP inhibits its activity, but under physiological conditions ATP inhibition is almost completely counteracted by FBP (Wang et al., 2001). Additionally, PK is regulated by developmental and tissue-specific mechanisms. In humans four tissue-specific PK isoenzymes have been identified: M1 (muscle), M2 (fetal and tumoral proliferating tissues), R (erythroid) and L (liver) (Fothergill-Gilmore and Michels, 1993). M1PK is expressed in adult normal tissues, like skeletal muscle, brain and heart, while PKM2 isoform is typically expressed in proliferating tissues (fetal, tumoral and several other adult tissues) (Guguen-Guillouzo et al., 1977). L-type PK (LPK) is predominantly expressed in the liver, but also in renal cortex, small intestine and pancreatic beta cells (Takegawa et al., 1983), while R-type PK (RPK) is specifically expressed in red blood cells (RBCs). During development, M2 is progressively replaced by the other tissue specific isoforms, although it is still the major isoform in leukocytes, platelets, spleen, kidney, lung and adipose tissue. Two structural genes (*PKM* and *PKLR*) encode these four PK isoforms. In humans M1 and M2 isoforms are encoded by the *PKM* gene, which is located in the q22 region of chromosome 15 (Takenaka et al., 1991), and are expressed from a single mRNA processed by alternative splicing. RPK and LPK isoforms are both encoded by the *PKLR* gene

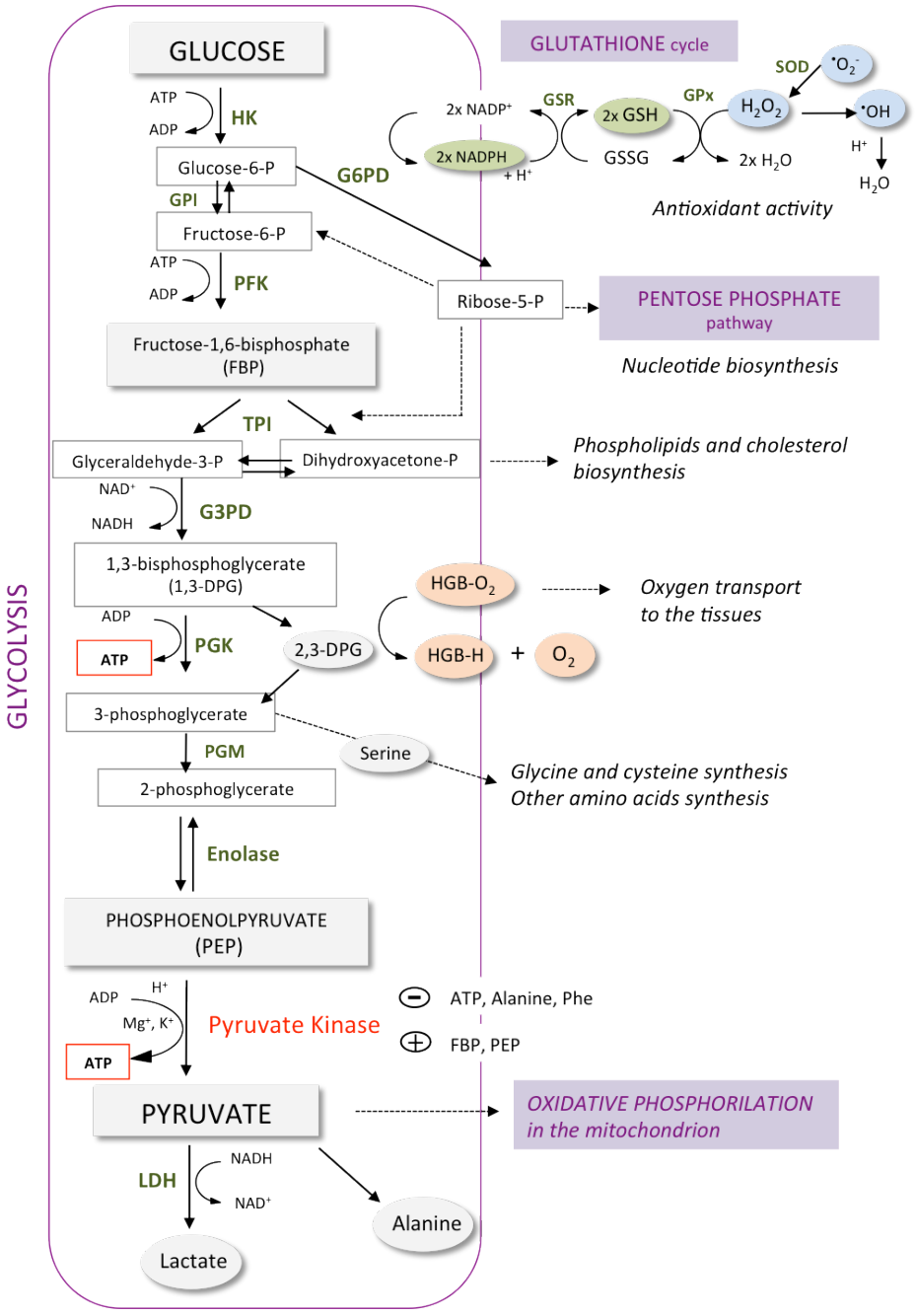

Figure 1: Schematic representation of glycolysis pathway and metabolic energy production indicating the connection with different biosynthetic processes. In the cytoplasm of the cells, glucose is metabolized to pyruvate via glycolysis, yielding two ATP molecules per molecule of glucose. PK catalyzes one of the two ATP-producing reactions of the glycolytic pathway and the metabolic intermediates of the route are connected with important anabolic processes in the cell, contributing to cell homeostasis and survival (adapted from (Lunt and Vander Heiden, 2011). PK is allosterically activated by FBP and displays a positive cooperativity towards PEP. On the contrary, ATP, alanine and phenylalanine (Phe) act as negative regulators of the enzyme. HK: hexokinase; G6PD: glucose-6phosphate dehydrogenase; GSH: reduced glutathione; GSSG: oxidized glutathione; GSR: glutathione reductase; GPx: glutathione peroxidase; SOD: superoxide dismutase; GPI: glucose-phosphate isomerase; PFK: phosphofructokinase; TPI: triosephosphate isomerase 2,3-DPG: 2,3-diphosphoglycerate; PGK: phosphoglyceratekinase; PGM: phosphoglycerate mutase; LDH: lactate dehydrogenase.

(chromosome 1q21) (Satoh et al., 1988) and their expression is regulated by the use of tissue-specific alternative promoters (Noguchi et al., 1987). In mice M1 and M2 PKs are encoded by the *Pk-3* gene located on chromosome 9 (Fitton and Bulfield, 1989), while the *Pk-1* gene encodes LPK and RPK isoforms and is located on chromosome 3 (Nijhof et al., 1984). Both are regulated by the same mechanism as that found in humans. The differential PK expression implies that each isoenzyme will exhibit different regulatory mechanisms and kinetic properties, adapting the enzymatic activity to the particular metabolic requirements of the expressing tissues.

PK has been largely conserved through evolution, and normally is a 200-240 kDa homotetramer comprising four identical subunits in almost all organisms (Zanella et al., 2005). RPK from human mature RBCs has been reported as a heterotetramer (L'_2Lc_2) built up by two subunits of 62 kDa (full-length subunits; 2L') and two of 57-58 kDa (subunits 2Lc), the latter resulting from a proteolytic cut at their N-terminus (Figure 2) (Kahn and Marie, 1982). This conformation enables the regulation of the enzymatic activity by switching from a low-affinity state of the PK protein for PEP (T), to a high-affinity state (R). In fact, the proteolysis of the N-terminal fragment has been suggested to play a role in PK regulation in RBCs by affecting the allosteric properties of the enzyme, as it participates in the intersubunit and interdomain interactions (Rigden et al., 1999).

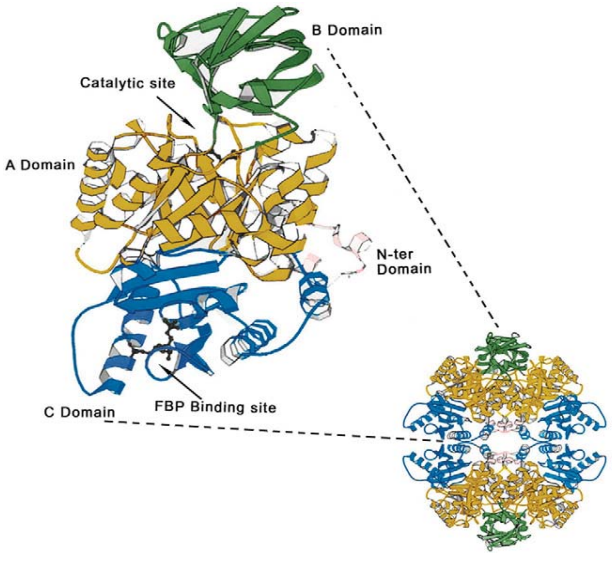

Figure 2: Three dimensional crystal structure of the RPK monomer (upper left) and tetramer (low right). RPK is a heterotetramer protein in which each subunit is organized into four domains: a small N-terminal helical domain; an A domain with $(\beta/\alpha)_8$ topology; a β-stranded B domain; and a C domain with α+β topology. The active site resides between the A and B domains, whereas the allosteric site is located in a pocket of the C domain. The four subunits of the tetramer are assembled to form a D2 symmetric oligomer. Both, A and C domains are involved in intersubunit interactions. (From Zanella et al. (Zanella et al., 2007)).

2. Metabolic aspects and cellular pathology in PK deficient erythrocytes

Erythrocytes function to transport oxygen from the lungs to the tissues via hemoglobin (HGB) and they are also important for maintaining the pH and osmotic balance in the blood via carbonic anhydrase. Most of mature RBC metabolic needs are covered by glycolysis, the oxidative pentose phosphate pathway and the glutathione cycle, but the energy needed for maintaining the cell membrane charge and for cycling nicotinamide adenine dinucleotide (NAD^+) is generated nearly exclusively through glycolysis (Wang et al., 2001). In addition, because RBCs loose the mitochondria, the nucleus and the endoplasmic reticulum during their differentiation process, they are unable to synthetize proteins *de novo* or to metabolize pyruvate via oxidative phosphorylation, relying almost exclusively on glycolysis for producing ATP and maintaining cell integrity. Despite its low efficiency (yielding only 2 ATP molecules per glucose), glycolysis can increase the ATP production by increasing the PEP catalysis rate (Pfeiffer et al., 2001), displaying a central role in RBC metabolism. Therefore, any loss of RPK activity will impair RBC metabolism and lifespan.

RPK is quite a stable protein to last the entire life span of erythrocytes (Wang et al., 2001). Reduced levels of RPK lead to two major metabolic problems: ATP depletion and increased 2,3-DPG content, affecting RBC homeostasis and survival (Lunt and Vander Heiden, 2011). The precise mechanisms that lead to a shortened lifespan of PK deficient RBCs are currently unknown, but metabolic block due to glycolytic intermediates accumulation has been proposed (Zanella et al., 2005). Additionally, low levels of intracellular ATP have been associated with hemolysis and membrane rigidity by dehydration via potassium and water loss (Miwa and Fujii, 1985), which can finally lead to apoptosis (Tsujimoto, 1997). The low ATP level can promote both K^+ efflux and Ca^{2+} accumulation in PK deficient RBCs, which respond with a reduction of energy consuming processes, such as cation pumps and phosphorylation of membrane proteins. In fact, the increased intracellular Ca^{2+} levels in PK deficient RBCs has been seen to promote the degradation of Band 3 membrane protein via μ-calpain activation (Passow, 1986), which likely alters the structure and properties of the erythrocyte membrane, leading to a shortened red cell lifespan (Kay, 1992). On the other hand, in some pyruvate kinase deficiency (PKD) patients, 2-fold and 4-6-fold increment in 2,3-DPG and glucose-6-phosphate

levels respectively have been observed (Lakomek et al., 1994), which can further impair the glycolytic flux by inhibition of hexokinase (HX) (Zanella and Bianchi, 2000). A further complication that PK deficient RBCs have to face is the reduced concentration of reduced nicotinamide adenine dinucleotide-phosphate (NADPH) (Zerez and Tanaka, 1987), which is essential for maintaining cellular redox balance and has an important role in protecting cells from oxidative damage. Reduced glutathione is indeed indispensable for maintaining normal erythrocyte structure. This ionic and metabolic imbalance caused by severe alterations in RPK dramatically affects the functionality of deficient RBCs, leading to intravascular hemolysis and their premature destruction in the spleen and liver.

3. Erythropoiesis

In humans, RBC life span in the blood stream is approximately 120 days. Erythroid cells are continuously produced in the fetal liver and in the adult bone marrow (BM) from a small population of pluripotent hematopoietic stem cells (HSCs) by the erythropoiesis process, which lasts around 5 to 6 days, to maintain hematopoietic homeostasis. The earliest erythroid committed progenitors identified *ex vivo* are the slow proliferating burst-forming unit-erythroid (BFU-E), which divide and differentiate into the rapidly dividing colony-forming unit-erythroid (CFU-E). The CFU-Es undergo erythroid differentiation through changes that include a decrease in cell size, chromatin condensation and hemoglobinization, leading up to extrusion of the organelles and enucleation (Figure 3). Erythropoiesis takes place in the erythroblastic islands of the BM (Figure 3.C), composed by a central macrophage that is thought to be a central regulator in the process, surrounded by differentiating erythroid cells (Walkley, 2011).

Erythropoietin (Epo) is the major physiological cytokine that promotes erythropoiesis under basal and stress conditions, and the extracellular matrix protein fibronectin is also important (Eshghi et al., 2007). Epo receptor (EpoR) is expressed by the earliest erythroid progenitors at the BFU-E stage (Wu et al., 1995), and its expression is lost as erythroid progenitors undergo terminal differentiation. Epo production in the kidney is induced by hypoxia and stimulates the proliferation and differentiation of CFU-E progenitors, which are highly Epo dependant. Binding Epo to EpoR on the surface of

erythroid progenitors triggers the activation of multiple intracellular pathways, including Jak2/STAT5, PI3K/AKT and Ras/MAPK signalling pathways (Richmond et al., 2005), which lead to the activation of an erythroid-specific genetic program.

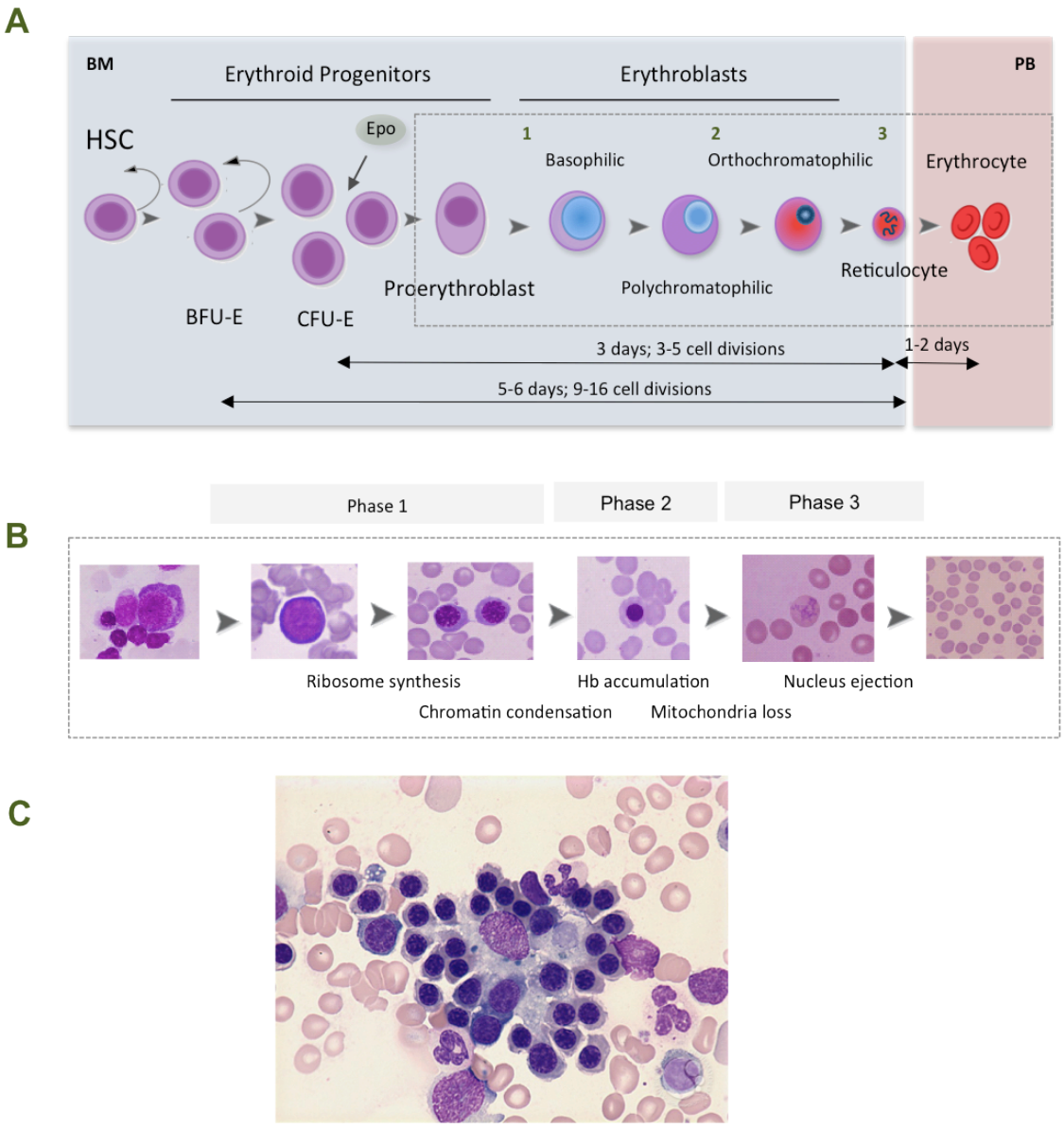

Figure 3: Erythropoiesis overview. **A)** The process by which multipotential HSCs differentiate into mature, non-nucleated erythrocytes is shown. The requirement of Epo starts between BFU-E and CFU-E stages, as the cell begins to express EpoR, and diminishes with erythroblast differentiation. In humans this process lasts 5-6 days in the bone marrow, and the reticulocyte leaves the erythroblastic island and enters circulation once it has lost the nucleus, which lasts 1 or 2 days. **B)** Photographs of the different stages of erythroid differentiation from the committed proerythroblasts progenitors to mature RBCs. This process takes place in three main phases entailing ribosome synthesis, chromatin condensation, Hb accumulation, loss of mitochondria and nucleus ejection. **C)** Developing red cell precursors from an erythroblastic island (picture from http://ashimagebank.qc.astutetech.com, author Peter Maslak).

In erythroid cells, STAT5 and other Epo-regulated transcription factors interact with several lineage-restricted transcriptional regulators, such as GATA-1, SCL/Tal1, LMO2, LDB1, Klf1 and Gfi-1b (Kerenyi and Orkin, 2010). Transferrin receptor (Tfr1) is also important during erythropoiesis, as is required for erythrocytes to uptake transferrin-bound iron (Rivella, 2012). Tfr1 expression is activated by STAT5 (Kerenyi et al., 2008) during the first divisions of CFU-E progenitor cells before entering into the erythroid differentiation (Hattangadi et al., 2011), and decreases as red cells mature (Ney, 2011). In addition, chromatin epigenetic modifications, like methylation of H3K479me2 and H3K4 histones, and miRNAs regulation (miR150 and miR223) also contribute to the tight regulation of the different stages in erythropoiesis (Hattangadi et al., 2011).

BFU-E progenitors respond to Epo, stem cell factor (SCF), insulin-like growth factor (IGF-1), glucocorticoids (GCs) and interleukines (IL3 and IL6). These and other cytokines regulate divisions of BFU-E cells and control their self-renewal and differentiation into more mature CFU-E progenitors (Hattangadi et al., 2011). Although erythropoiesis regulation is not fully understood, it is known that the hematopoietic microenvironment regulates the action of Epo on CFU-E, determining the number of CFU-Cs available to start the differentiation into proerythroblasts (Walkley et al., 2011).

Proerythroblasts are big cells and the nucleus occupies 80% of the cytoplasm, in which the polyribosomes are abundant and ferritin and hemoglobin are also present. In the next step of differentiation (basophilic erythroblast) the size of the nucleus is progressively reduced and the chromatin starts to condense, but the polyribosomes are still present. These erythroid progenitors differentiate into a smaller polycromatophilic erythroblast, in which the nucleus occupies less than half of the cell volume and the heterochromatin is condensed. In this state hemoglobin overtakes the polyribosome content and the rest of organelles start to disappear. The last mitotic division of the erythroid lineage gives rise to the orthochromatophilic erythroblasts, the smallest progenitor in the series, in which the hemoglobin concentration is very high and the nucleus occupies a quarter of the cell volume in a non-concentric position. In this state, the mitochondria reduce their size and the ribosomes are dispersed. The last step of erythroid differentiation within the BM produces reticulocytes that still have some mitochondria and a residual nucleus. The nucleus ejection takes place in the erythroblastic island or while reticulocytes leave their niche in the BM through the sinusoids and enter circulation (Figure 3).

Pyruvate kinase shows a differential expression throughout the erythroid maturation. Mature RBCs express the RPK isoform and in immature erythroid progenitors the PKM2 isoform is predominant (Max-Audit et al., 1984). During the erythropoiesis process, PKM2 expression decreases gradually, and the RPK synthesis increases simultaneously through the activation of the PKLR promoter by erythroid-specific TFs (Takegawa et al., 1983). CFU-E progenitors express both isoforms and in proerythroblasts the PKM2 expression is still predominant. PKM2 expression decreases progressively being undetectable in orthochromatophilic erythroblasts. Accordingly, RPK expression is higher in basophilic and orthochromatophilic erythroblast than in proerythroblasts (Nijhof et al., 1984).

Under normal conditions, approximately 1% of erythrocytes are synthesized each day, but RBC production can substantially increase during times of acute or chronic hemolysis. In mice, the normal distribution of murine erythroid progenitors is 90% in the BM and 10% in the spleen (Richmond et al., 2005), but the spleen can expand up to tenfold during chronic erythroid stress. In this anemia-induced situation, the number of CFU-E progenitors is insufficient to produce the required RBCs even under high Epo levels, and the body responds by producing more CFU-Es from BFU-Es (Hattangadi et al., 2011). In humans erythropoiesis occurs only in BM, and splenomegaly is considered as a consequence of the reticulocyte sequestration rather than extramedullar erythropoiesis (Aizawa et al., 2005). However, under atypical conditions such as acute anemia, spleen and liver can become erythropoietic sites. In fact, spleen extramedullar hematopoiesis has also been observed in some PKD patients together with an increased apoptosis in the erythroid compartment of the spleen, which led to ineffective erythropoiesis caused by *PKLR* mutations (Aizawa et al., 2003).

Acute anemia induces a physiological response that implies the rapid development of new erythrocytes in a process called stress erythropoiesis. Studies from Wang et al proposed that Epo was the main regulator of this response, increasing its synthesis in the kidney as a result of the tissue hypoxia, and driving the expansion and differentiation of BM progenitors (Wang and Semenza, 1996). Eventually, some of these progenitors would then migrate to the spleen and finish their differentiation.

4. PKLR GENE ORGANIZATION AND TRANSCRIPTIONAL REGULATION

PKLR gene was first discovered in 1961 (Valentine et al., 1961) and covers a sequence of around 9.5 Kb. The coding region is split into 12 exons, 10 of which are shared by the two isoforms, while exons 1(R) and 2(L) are specific for RPK and LPK respectively. The transcriptional control depends on the regulatory activity of the promoter sequences flanking each first exon (R or L) allowing its differential expression. Transcriptional regulation of LPK expression in liver is also controlled by hormonal and glucose stimulation (Decaux et al., 1989). Human RPK cDNA is 2060 bp long and encodes a protein of 574 amino acids (Kanno et al., 1991), whereas LPK cDNA contains 1629 bp encoding a 543 amino acids monomer (Figure 4).

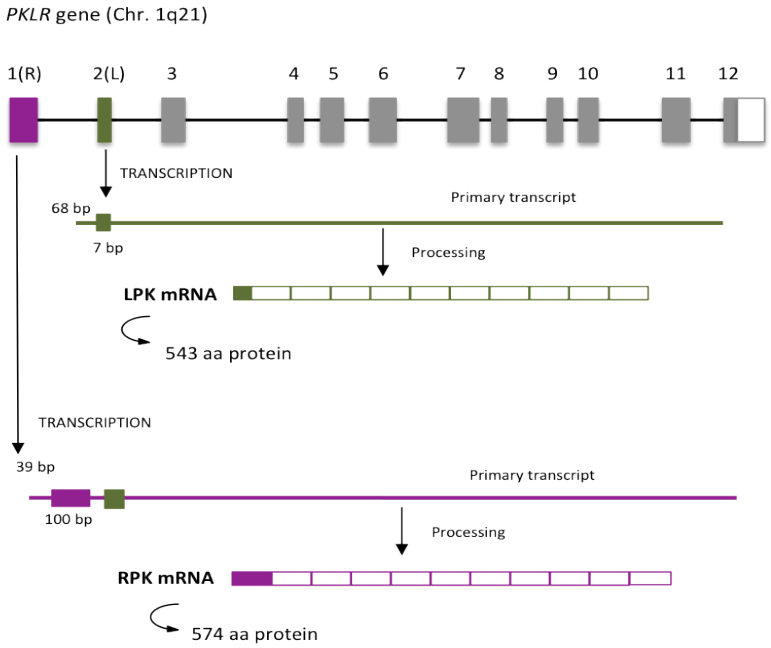

Figure 4: Structure of human *PKLR* gene and schematic representation of RPK and LPK isoforms expression. Human *PKLR* gene is located on chromosome 1q21 and covers around 9.5 kb. Exons R and L are indicated by violet and green boxes respectively. The exons common to both mRNAs are shown in grey. The white box represents the 3' non-translated sequence of exon 12. LPK cDNA contains 1629 bp encoding a 543 amino acid monomer, and transcripts contain 68 bp from the 5' non-translated region and 7 bp of the coding region (2 amino acids). RPK cDNA is 2060 bp long and encodes a 574 amino acids monomer. Transcripts include 39 bp of the 5' untranslated region and the first 100 bp of the coding sequence (33 amino acids), yielding a polypeptide sequence of 574 amino acids, being 31 amino acids longer than LPK. Modified from (Kanno et al., 1992) and (Noguchi et al., 1987). *Notice that some authors name exon 1 as 1(R) and exon 2 as 1(L), modifying the numbering of the rest of the exons*.

RPK polypeptide sequence is 31 amino acids longer than the liver-specific PK because of the use of alternative first exons in each specific-tissue isoform. LPK is expressed when transcription starts in exon 2(L), containing 68 bp from the 5' non-translated region and 7 bp of the coding region (2 amino acids). In erythroid cells, however, transcription starts in exon 1(R) which covers 39 bp of the 5' untranslated region and the 100 bp coding sequence (33 amino acids) (Kanno et al., 1991). The splice acceptor site of intron 1 does not have any consensus splicing sequences, so it is thought that different processing of RPK primary transcript will remove exon 2 and intron 2, as well as intron 1 (Figure 4).

The R-type promoter region of the *PKLR* gene has been located in the 5' flanking sequence upstream of the first exon (1R) (Kanno et al., 1992). This region covers the first 270 bp upstream of the translation initiation codon and has neither typical TATA or CCAAT boxes. The PKLR promoter contains two CAC boxes (binding members of the CCACC/Sp family of TFs involved in gene expression during early developmental stages), as well as four GATA motifs and the PKR-RE1 element (van Wijk et al., 2003), whose mutation has been involved in severe cases of PKD (Figure 5).

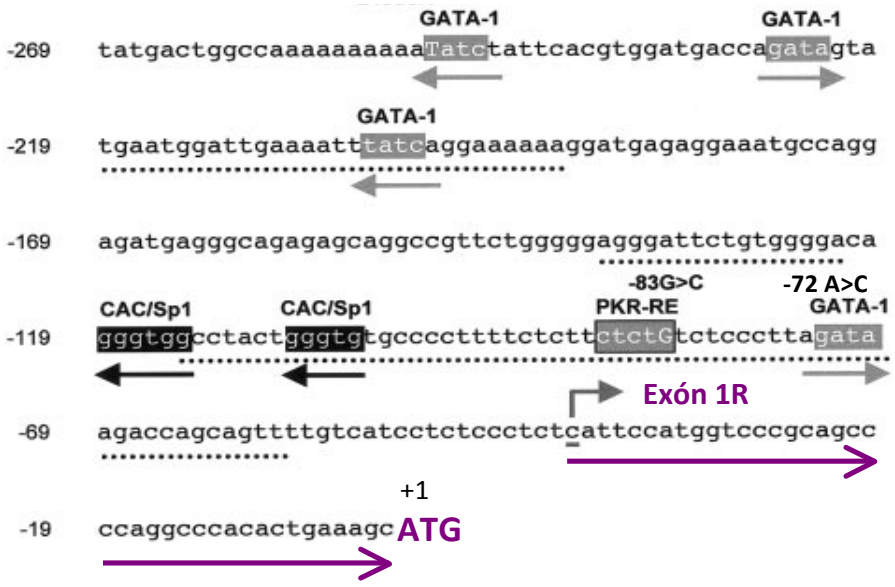

Figure 5: DNA sequence of the human erythroid-specific PKLR promoter. A 269 bp region comprising the upstream regulatory domains and exon 1 upstream of the ATG start codon is shown. GATA-1, CAC/Sp1 and the regulatory PKR-RE1 element are shown in boxes, and their orientation is indicated by an arrow. The two *cis* mutations identified in the PKLR promoter are also indicated in the sequence: -83 G>C affecting the PKR-RE element (van Wijk et al., 2003) and -72 A>G mutation involving the GATA-1 element, both causing severe hemolytic anemia (Manco et al., 2000). Dashed line indicates conserved elements between rat and human PKLR promoter (Kanno et al., 1992). Modified from (van Wijk et al., 2003).

Studies from Kanno et al have shown that the proximal 120 bp region has a basal promoter activity, while distal 150 bp region from the ATG codon, acts as a powerful enhancer in erythroid cells (Kanno et al., 1992). Moreover, studies from Max-audit et al have demonstrated that binding of the erythroid transcription factor GATA-1 to the GATA proximal element (-69 nucleotides upstream from the first codon) is required for RPK minimal promoter activity, and it is also a potential TATA box (TAGATAA) (Max-Audit et al., 1993) (Figure 5). As already demonstrated for the human β-globin gene, combinations of these two motifs (CAC/SP1 and GATA) cooperatively direct the erythroid specificity of β-globin promoter. Nevertheless, their association has been seen in positive, negative and inducible regulatory sequences, which suggests that other elements may be contributing to the fine-tuning of erythroid gene expression (Raich and Romeo, 1993).

5. Genetic characteristics of human PK deficiency (PKD)

PKD is the most frequent hereditary enzymatic defect of the glycolysis pathway in RBCs causing chronic non-spherocytic haemolytic anaemia (CNSHA), and there are more than 400 cases described so far. The large number and variety of *PKLR* mutations reported to cause PK deficiency suggest that the RPK protein is highly sensitive to even small changes, affecting protein folding, thermostability, catalytic efficiency or allosteric response to effectors. This can have serious consequences for long-term erythrocyte function, as mature cells lack the machinery to compensate PK deficiency by generating more protein.

PKD is transmitted as an autosomal recessive trait and clinical symptoms usually occur in homozygotes or compound heterozygotes for two mutant alleles. Around 195 different mutations in the *PKLR* gene are associated with CNSHA and 8 polymorphic sites have been reported so far, located throughout the whole coding region with no preference for the active site. PKD disease has a worldwide distribution and three of the most common mutations (1529A, 1456T and 1468T) have a strong ethnical and regional background (Table 1). 1529A is the most common in the USA (42%) and in Northern and Central Europe (41%); 1456T is most common in Southern Europe (32% in Spain, 35% in Portugal and 29% in Italy), where 1529A is rare; and 1468T occurs more frequently in Asia (Zanella et al., 2005). All

these mutations were found in the context of its own haplotype, suggesting that each has a unique origin. Interestingly, strong evidence suggests a protective role of some PK deficient variants in malaria endemic areas (Berghout et al., 2012); (Machado et al., 2010) and (Machado et al., 2012).

Prevalence of PKD varies considerably, but it has been estimated at 1:20 000 in the general Caucasian population (Beutler and Gelbart, 2000). The highest prevalence has been found in Saudi Arabia (3.12% in newborns; (Abu-Melha et al., 1991) whilst the lowest frequency of PKD was reported in Hong Kong Chinese population (<0.21%; (Fung et al., 1969). However, PKD prevalence is thought to be underestimated because of the common RPK activity overlap between heterozygous patients and normal values, and sometimes between those and homozygous patients (Beutler and Gelbart, 2000). PKD prevalence is thought to be higher due to the low survival rate of severely affected patients, leading to underestimate. It is also expected to be higher in communities where consanguineous relationships are common (i.e *Gypsy* deletion, Table1, or PK *Amish* variant, Table 2).

Most PKD patients show missense (69%), splicing (13%) and nonsense (5%) mutations, while small deletions, insertions and frameshift mutations are rare (Zanella et al., 2007). The severity of the syndrome is commonly associated with mutations involving the active site, the allosteric binding site or the protein stability, usually located in highly conserved amino acid residues (Zanella et al., 2005). Only two mutations, -72 A>G and -83 G>C (Figure 5), have been identified in the promoter region and functionally characterized so far, both associated with severe PKD phenotype. The first one disrupts the GATA-1 motif (-69 to -74) (Manco et al., 2000) and the second one alters the regulatory element PKR-RE1 (-87 to -83) (van Wijk et al., 2003). At the moment, only four large deletions have been identified: the "*Gipsy*" deletion of 1149 bp losing exon 11 (Baronciani and Beutler, 1995), the PK "Viet" deletion losing exons 4 to 10 (Costa et al., 2005), a deletion of 5006 bp lacking exons 4 to 11 (Fermo et al., 2005) and a 1477 bp deletion (c1618+37_2064del) removing most of intron 11, the coding sequence of exon 12 an almost half of the 3' untranslated region (van Wijk et al., 2009) (Table 1).

cDNA nucleotide substitution	Effect	Exon(s)	Distribution		Reference
			Ethnical	Geographical	
- 83 G>C		Promoter			van Wijk et al., 2003
-72 A>G		Promoter			Manco et al., 2000
5006 bp (***IVS3 > nt 1431***) del	ex 4-11 del	IVS3-ex10			Fermo et al., 2005
ex 4-10 del (***PK Viet***)	ex 4-10 del	4-10	Asian	Vietnam	Costa et al., 2005
721 G>T	241 Glu > End	7	low freq in Caucasian	Asia	Baronciani and Beutler 1993
G829 G>A	277 Glu > Lys	7	Sub- sahanrian, Mandeka and Brahui	West Africa and Pakistan	Berghout et al., 2012 Machado et al., 2012
994 G>A	332 Gly >Ser	8	low freq in Caucasian	Asia	Lenzner et al., 1994
1151 C>T (***Tokyo***)	384 Thr > Met	9	Japanese and Caucasian		Neubauer et al., 1991
1436 G>A	479 Arg > His and sp.site	10	Japanese and Caucasian		Kanno et al., 1994a
1437-1618 del (***Gypsy***)	Frameshift	11	Gypsy Caucasian		Baronciani and Beutler 1995
1456 C>T	486 Arg > Trp	11	Caucasian	Sothern Europe	Baronciani and Beutler 1993
1468 C>T	490 Arg > Trp	11	Asian	Asia	Kanno et al., 1994c
1529 G>A	510 Arg > Gln	11	Caucasian	North America and Northern Europe	Baronciani and Beutler 1993 Lenzner et al., 1994
c1618+37_2064del 1477 bp	552-end	I11-ex12	Caucasian	Northern Africa	van Wijk et al., 2008

Table 1: Some of described *PKLR* mutations causing CNSHA. The table shows the *PKLR* sequence affected by the mutation, as well as the ethnical and geographical distribution of these mutations. Three of the most common mutations are highlighted in green.

6. CLINICAL ASPECTS OF HUMAN PKD

6.1 PKD clinical features

The clinical phenotype of PKD is heterogeneous because of the wide number of mutations involved, and its severity depends on the type of sequence that has been modified and its location relative to the initiation codon. Heterozygous patients do not show clinical or hematological symptoms of hemolysis because they still have residual RPK activity (approximately 50%) (van Wijk et al., 2009), and the majority of reported PKD patients are homozygotes and compound heterozygotes carrying two different mutations in the *PKLR* gene. In these patients the degree of hemolysis varies considerably because they have different hybrid PK tetramers with distinct structural and kinetic properties. As shown in table 2, carriers of two different mutations can also show a very severe PKD phenotype.

The *PKLR* gene uses alternative first exons transcription as regulatory mechanism, so defects in *PKLR* gene may result in alterations of both erythroid and liver enzymes. However, clinical symptoms of PKD are confined to RBCs because the hepatic deficiency is usually compensated by the persistent synthesis of the LPK enzyme in hepatocytes (Nakashima et al., 1977). Pathological manifestations are usually observed in patients with less than 25% of the normal RPK activity suffering lifelong chronic hemolytic anemia and splenomegaly (Lenzner et al., 1997); (Zanella et al., 1997); (Zanella et al., 2001a). In PKD patients, splenomegaly is attributed the intense sequestration of damage RBCs, but also as a consequence of the extramedullar erythropoiesis (Aizawa et al., 2003).

The degree of anemia in PKD patients varies widely, ranging from mild or fully compensated asymptomatic anemia, to severe hemolytic neonatal anemia with pronounced jaundice, necessitating blood transfusions and continuous therapy. In addition, hematological features of PKD are common to other diseases causing CNHSA, with hemoglobin levels of 9.8 g/dL in PKD unsplectomized patients and 7.3 g/dL in splenectomized ones. Extreme reticulocytosis (166.10^9 retics/L or 4.66%) and the presence of erythroblasts in peripheral blood are common clinical signs of severe PKD, especially after splenectomy (Zanella and Bianchi, 2000); (Sedano et al., 2004). However, in humans

Patient	Mutation	Effect	Country	Symptoms	Clinical history	Reference
Homozygote (1 patient)	null mutation 1133-11nt ins exon 8	frameshift truncated protein with 6 aa more altered active site	Spain	low weight at birth, low Hb jaundice, oedema hepatosplenomegaly First blood transfusion at birth severe transfusion-dependant hemolytic anemia	Unknown etiology hydrops foetalis Asymptomatic heterozygous parents each with 50% normal RPK activity, kinetics and stability Splenectomy at 3 years old Death at 11 years old because of allogenic BMT complications	Gilsanz et al., 1993 Diez et al., 2005
Compound Heterozygote (2 patients)	Paternal Allele: 948 C>G (***PK Saigon***) Maternal Allele: ex4-10 del (***PK Viet***)	Exon 7 del ex4-10 reduced RPK activity	South-east Asia	Low Hb, jaundice hemolytic anemia hepatosplenomegaly Severe iron overload First blood transfusion at 3 days-old severe transfusion-dependant	Heterozygous parents for PKD with blood transfusions history Maternal grandmother β-thalassaemia not involved in PKD outcome	Costa et al., 2005
Homozygote (7 patients)	nonsense mutation 1318 G>T (***PK Aarau***)	exon 10 truncated protein with 33 aa less	Switzerland	low weight at birth, low Hb jaundice Splenomegaly severe transfusion-dependant hemolytic anemia	large antecedents of anemia high consanguinity splenectomy at 2-10 years old death of 5/7 patients at 2 days after birth	Sedano et al., 2004
Compound Heterozygote (1 patient)	Paternal Allele: del 5006 bp (IVS3-nt 1431) Maternal Allele: missense 409 G>A	del ex4-11 exon 5	Australia	severe neonatal anemia First exchange transfusion at birth	Recurrent abortoins at first trimester of gestation Death soon after birth	Fermo et al., 2005
Compound Heterozygote (1 patient)	Paternal Allele: 862 G>T Maternal Allele: 1436 G>A (***PK Amish***)	exon 7 altered PEP binding site	Japan	severe neonatal jaundice first blood transfusion at 3 months old Fototherapy severe transfusion-dependant hemolytic anemia	Heterozygous parents for PKD Splenectomy at 4 years old	Aizwa et al., 2003

Table 2: Severe PKD phenotype. Table summarizes some of the severe PKD cases reported worldwide, indicating the genetic defect in the *PKLR* gene, the symptoms and the clinical history.

reticulocytosis is not proportional to the severity of hemolysis because young PKD erythrocytes are selectively sequestered in the spleen, and after splenectomy the levels dramatically increase even when the anemia becomes less severe.

Gallstones and iron overload are common problems secondary to hemolytic anemia, and are usually caused by chronic transfusion therapy (Zanella et al., 2005), but it may even occur in non-transfused patients dramatically increasing morbidity in PKD patients. The cause of iron overload in PKD pathology is considered multifactorial (Zanella et al., 2001b), and both splenectomy and ineffective erythropoiesis have been proposed as cofactors (Zanella et al., 1993). Thus, iron overload is a serious problem caused by hemolytic anemia, as shown in four homozygous PKD patients (1529A), reported as untransfused and splenectomised, that developed liver cirrhosis and cardiomyopathy secondary to iron overload in the absence of hereditary hemochromatosis (HFE) gene mutation (Zanella et al., 1997). Two of them died because of hemochromatosis complications.

Unconjugated bilirubin concentration is very often increased, but usually < 85 μmol/dL), and may show a slight rise after splenectomy. Red blood morphology observed in PKD blood smears is not specific, and both osmotic fragility and autohemolysis are only present in 20-25% of studied patients (Zanella and Bianchi, 2000). Growth retardation, *hydrops foetalis* and death during the neonatal period have also been reported in severely affected patients (Gilsanz et al., 1993), (Sedano et al., 2004), (Ferreira et al., 2000) (Table 2). In addition, as mentioned before, enhanced extramedullar hematopoiesis and increased apoptosis of erythroid cells have been observed in the spleen of a PKD patient (Aizawa et al., 2003) and in mouse models (Aizawa et al., 2005).

6.2 Compensatory mechanism in PKD: PKM2 overexpression and ineffective erythropoiesis

The clinical severity of PKD has also been associated with a compensatory persistence of the PKM2 isoform expression in patient RBCs (Lenzner et al., 1997) y (Kanno et al., 1994). During the physiological erythropoiesis process, PK expression switches from PKM2 in erythroblasts to the RPK specific isoform in mature RBCs in response to developmental signals (Max-Audit et al., 1988). Compensatory PKM2 expression is

attributed to its overproduction in erythroblasts, analogous to the compensatory fetal hemoglobin production in thalassemia major (Miwa et al., 1975). PKM2 expression could cover the energetic needs of deficient mature RBCs, partially explaining the survival of severe PKD patients. However, it has also been reported that patients with complete deficiency of the RPK isoform survive in the absence of the PKM2 isoenzyme in RBCs (Diez et al., 2005), suggesting that other mechanisms may be acting during PKD pathogenesis, such as glycolytic signals or defective RPK isoforms overexpression.

Deficient RBCs show a shortened lifespan caused by ATP depletion, and they are quickly removed from the blood by the reticuloendothelial system. Most PKD patients respond to this lack of mature RBCs increasing the number of circulating immature cells (erythroblasts and reticulocytes), because these cells already have some metabolic capabilities (Keitt, 1966), cover the patient's minimal metabolic demands (Lakomek et al., 1989). Simultaneously, PKD leads to an increase in extramedullar hematopoiesis that in severe cases fails to compensate anemia.

6.3 PKD diagnosis

PKD diagnosis is not routinely performed, and is normally based on determination of RPK activity and DNA sequencing because the hematological symptoms are not distinctive enough for this disease. Neither the residual enzymatic activity nor the reticulocyte number is related to the severity of the hemolysis in PKD patients (Zanella et al., 2005). Some PKD patients display an increased RPK activity attributed to the higher RPK activity in reticulocytes than in mature RBCs, and other patients show an overexpression of aberrant RPK to compensate the anemia (Diez et al., 2005). Besides, general differences in energy metabolism of patients of different ages and under different treatments cause overlap between normal and heterozygous levels, complicating PKD diagnosis. Other pitfalls in the PKD enzymatic diagnosis are the low amount of protein obtained for testing, the presence of hybrid enzymes in compound heterozygotes, sample contamination with PKM2 isoenzyme from leukocytes and the presence of healthy erythrocytes in transfused patients, together with the compensatory PKM2 expression in RBCs observed in some PKD severe patients (Lenzner et al., 1997) (Kanno et al., 1994). On the other hand, prenatal testing of PKD is

required in cases with severely affected siblings, helping in managing acute symptoms in newborns (Baronciani and Beutler, 1994).

During the last twenty years, important efforts have been focused on the production and characterization of recombinant mutant proteins of human RPK. Through structural three-dimensional analysis, a correlation between the function and location of the mutated amino acid and the type of molecular defect has been found in a large number of *PKLR* mutations. However, an association has not been proven between the residual enzymatic activity and clinical severity of the disease (Valentini et al., 2002). In general, most authors agree that PKD clinical manifestations reflect the complex interactions of additional physiological and environmental factors, such as genetic background, concomitant polymorphisms of other enzymes, infections, posttranscriptional or epigenetic modifications, differences in intracellular proteolytic activity in each patient, ineffective erythropoiesis and differences in splenic function (Zanella et al., 2005).

6.4 Conventional treatment for PKD

PKD treatment is based on supportive measures because no specific therapy for PKD is currently available. Some drugs and chemicals have been administered to improve PKD clinical symptoms (e.g AMP, $Mg2^{+}$, fructose), but none of them have been seen to be effective. Severe patients usually require blood transfusions (particularly at the first year of life), and sometimes exchange transfusions are needed soon after birth. Normally, these procedures considerably impair the quality of life and lead to iron overload problems. Interestingly, anemia can be well tolerated in some PKD patients because of increased 2,3-DPG in erythrocytes (Oski et al., 1971), which regulates the oxygen dissociation curve of hemoglobin.

Splenectomy can be clinically useful in some severe PKD patients because, although it does not arrest hemolysis, it often results in an increase of HGB level of 1-3 g/dL, avoiding transfusion requirement in most patients (Zanella et al., 2007). However, splenectomized patients are more susceptible to infections and thrombotic events, and aplastic or hemolytic crisis may still occur. As therapeutic efficacy of this procedure cannot be predicted, it is recommended only for severely affected, young patients who need regular blood

transfusions or patients who do not tolerate anemia (Zanella et al., 2005). Iron chelation therapy has been also used (i.e. desferroxamine, deferiprone) due to iron overload is a very common problem in PKD even in untransfused patients, (Zanella et al., 2001b); (Marshall et al., 2003), and Epo administration has been reported to be effective at reducing iron overload in one patient (Vukelja, 1994). Allogeneic bone marrow transplantation (BMT) has been used in specific severe cases, and has been successfully reported in one 5 year-old severely affected child (Tanphaichitr et al., 2000). The patient, showing compensatory erythropoiesis in BM and extramedullar hematopoiesis, was conditioned with busulfan and cyclophosphamide, and transplanted with BM cells from his HLA-identical sister, who had normal PK activity. After three years, normal HGB levels and RBC PK activity were observed, with no evidence of hemolysis. This case illustrates that BMT can be clinically successful by complete replacement of RPK deficient cells with normal RBCs. However, the toxicities from conditioning treatments and subsequent immunosuppression result in significant morbidity and mortality of transplanted patients.

7. PKD ANIMAL MODELS AND EXPERIMENTAL APPROACHES

Due to no effective treatment being available today for severe PKD patients, some experimental alternative therapies have been explored in animal models, which are rather useful tools for understanding the pathophysiology of PKD disease. Currently, several canine and murine strains have been characterized. All of them are naturally occuring mutants carrying defects in the *Pklr* gene, as no transgenic PKD strain has been yet developed.

7.1 PKD dog models

Among dog strains, the Basenji breed shows severe chronic hemolytic anemia because of a homozygous frameshift mutation (c.433delC deletion) causing a single base deletion (Whitney et al., 1994), and probably affecting the catalytic site of the RPK enzyme. Basenji dog RBCs display a shortened life span with compensatory overexpression of the PKM2 isoform (Black et al., 1978), and reticulocyte counts over 15%. They also show an

increased RPK activity, but with abnormal kinetic properties and impaired thermostability (Giger and Noble, 1991). Hemolytic anemia phenotype has been also described in West Highland White Terriers, which carry a 6 bp insertion in exon 10 affecting the C domain in RPK (Skelly et al., 1999), as well as in Beagle dogs showing 50% normal RPK activity and compensatory PKM2 expression in erythrocytes (Giger et al., 1991).

Experimental BMT assays performed in the Basenji dog model have shown that clinical benefits correlate with the degree of chimerism. The allogeneic BM transplantation of dog leukocyte antigen (DLA)-identical cells from healthy littermates' BM failed to resolve the hematological symptoms without previous recipient conditioning, but hemolysis was almost completely reverted when myeloablative conditioning was applied (Zaucha et al., 2001). Takatu et al found that around 20% of wild type donor repopulating cells were required to ameliorate the disease phenotype after non-myeloablative conditioning, but when they increased the level of donor chimerism by donor lymphocyte infusion (DLI), 47-62% of donor engraftment in the BM yielded a better clinical improvement (Takatu et al., 2003). Although this approach overcomes problems from high dose conditioning, DLI from non-sensitized donor could lead to aplasia and graft-versus-host disease (GvHD) (Chang and Huang, 2013).

7.2 PKD mouse models

In mouse, the first mutant strain associated with CNSHA was reported by Kanno et al in the CBA/N strain (CBA-*Pk-1*^{slc}/*Pk-1*^{slc}) carrying the 1013 G>A homozygous missense mutation in exon 8 of the *Pklr* gene (Kanno et al., 1995). The consequent amino acid substitution (338 Gly>Asp) has been suggested to cause a conformational change in the PEP binding site, analogous to the *PK Hong Kong* variant in humans (homozygous missense 941 T>C mutation or 314 Ile>Thr) (Kanno et al., 1994) in which the substitution involves an amino acid adjacent to the active site of RPK enzyme. CBA-*Pk-1*^{slc}/*Pk-1*^{slc} RPK deficient mice show around 90% reduction in PK activity, high splenomegaly degree (4.6 times more than wild type CBA mice), iron deposits in the spleen and a remarkable reticulocytosis rate (41.6%) (Morimoto et al., 1995). However, hemolytic anemia is less severe than expected in these mice when it is compared to the equivalent human phenotype, probably due to the

compensatory extramedullar hematopoiesis observed in the spleen from CBA-*Pk-1*^{slc}/*Pk-1*^{slc} mice.

This PKD CBA-*Pk-1*^{slc}/*Pk-1*^{slc} mouse model has been used to study the feasibility of BMT in PKD, resolving the hemolytic anemia after BMT without host irradiation (Morimoto et al., 1995). However, other studies have been conducted to selectively expand normal donor erythrocytes in minimally conditioned recipients by using chemical inducers of dimerization (CIDs). These small molecules mediate the activation of the intracellular portion of the thrombopoietin receptor in RBCs (Richard et al., 2000); (Jin et al., 2000), allowing the expansion of donor RBCs with normal RPK activity, and therefore improving the anemia in recipient mice with decreased Epo levels (Richard et al., 2004). This strategy could be useful for transplantation methods that require erythroid chimerism, although the risk of long-term drug exposure is still present.

Finally, the other identified mouse strain reproducing human PKD symptoms is the recombinant congenic AcB55 strain (Min-Oo et al., 2004), which was found during the screening of loci involved in malaria susceptibility (Fortin et al., 2001). These mice carry the 269 T>A loss-of-function mutation affecting exon 2 of the *Pklr* gene, which results in the amino acid substitution 90 Ile>Asn in the RPK enzyme, leading to malaria resistance of deficient RBCs (Min-Oo et al., 2003). The *Pklr* ^{269A} allele in AcB55 strain causes hemolytic anemia compensated by constitutive erythropoiesis, as well as splenomegaly and constitutive reticulocytosis, which in turn protects the mice against *Plasmodium chabaudi* infection (Min-Oo et al., 2004). Comparative studies have shown that 1013 G>A mutation in the CBA-*Pk-1*^{slc}/*Pk-1*^{slc} strain is more deleterious than 269 T>A mutation in AcB55 mice with respect to enzymatic activity and severity of hemolytic anemia, with a more dramatic reduction of the erythrocyte half-life (increased turnover) in CBA-*Pk-1*^{slc}/*Pk-1*^{slc} mice than in AcB55 (Min-Oo et al., 2007).

8. Innovative therapies for PKD

Conventional allogeneic BM transplantation allows the generation of donor-derived functional HSCs of all lineages in the host, and could serve as a therapeutic option for inherited diseases like severe PKD. However, the availability of human leukocyte antigen (HLA)-compatible donors limit the number of patients that can benefit from this treatment, and BMT-derived complications can be as severe as the protein deficiency. As PKD is a monogenic recessive inherited disease and the metabolic defect is restricted to the erythroid lineage, it can be a suitable disease to be treated by gene therapy. Indeed, its feasibility has been proven in several recent reports (Trobridge et al., 2012) (Meza et al., 2009).

Gene therapy is the treatment of a disease by the introduction of a therapeutic gene into appropriate cellular targets using integrative or non-integrative vectors. Genetic modification can be achieved by *in vivo* direct infusion of the vectors carrying the ectopic gene, or by *ex vivo* manipulation of patient's cells. In recessive diseases caused by the mutation of a single gene, the procedure will require 1) the isolation of HSCs/progenitors from the BM (or mobilized into peripheral blood), 2) their transduction with an integrative viral vector expressing a functional cDNA copy of the defective gene and 3) their infusion back into the patient. Genetically corrected HSCs will repopulate all hematopoietic lineages in the patient, and the stable expression of the transgene in the affected cells will replace its non-functional endogenous counterpart, restoring its function. Currently, HSC gene therapy is a therapeutic alternative for several non-curable diseases, such as monogenic diseases of the blood and immune system, as well as for lysosomal storage diseases.

In addition, HSC gene therapy has several advantages over the allogeneic BMT. Firstly, it overcomes the limitation of HLA-compatibility, so it can be applied to every patient. Secondly, no post-transplant immunosuppression is required as there is no risk of GvHD, reducing the morbidity and mortality of transplanted patients (Naldini, 2011). There is also a reduced risk of graft rejection because it is an autologous transplant, allowing the use of reduced-intensity conditioning regimens before the transplant. Gene therapy treatment following a reduced conditioning protocol has provided clinical benefits in the adenosine deaminase (ADA)- severe combined immunodeficiency (SCID) (Aiuti et al., 2002);

(Gaspar et al., 2006) and in Wiskott-Aldrich syndrome (WAS) clinical trials (Boztug et al., 2010).

8.1 Challenges in HSC gene therapy protocols

Integrative and non-integrative vectors can be used in gene therapy, and they can also be classified as viral or non-viral vectors. When long-term expression is needed, integrative vectors are required. In therapies based on HSC genetic modification, the exogenous therapeutic gene has to be specifically, efficiently and stably incorporated. However, some important issues have to be considered, which depend on the type of disease to be treated and on the vector design used to deliver the transgene to the target cells.

***Ex vivo* culture**. The genetic modification of HSCs requires the *ex vivo* manipulation of the cells, which can affect their self-renewal capability, viability, homing and repopulation capacity once they are transplanted *in vivo*. The development of new *ex vivo* transduction protocols using lentiviral (LV) vectors that require less cell manipulation and shorter culture times, has allowed the preservation of HSC biological features (Mostoslavsky et al., 2005), maintaining their repopulation properties in mouse models (Gonzalez-Murillo et al., 2008).

Transfer vector design. Integrative viral vectors have been used to achieve the stable integration of the therapeutic transgenes in the genome of the target cells. Among them, γ-retroviral (γ-RV) vectors have been traditionally used in clinical trials, but their integration requires cells to enter cell division, extending the time in culture and reducing transduction efficacy. LV vectors overcome these hurdles because they are able to reach the nucleus of non-dividing cells (Lewis and Emerman, 1994).

Gene transfer efficiency. Efficient gene transfer is essential and often represents the main limitation for developing preclinical protocols with therapeutic applications. In general efficiency improves with high viral titers. LV vectors have provided clinical benefits in several preclinical protocols in a transgene expression-dependent manner, such as β-thalassaemia (May et al., 2000), adrenoleukodystrophy (ADL) (Cartier et al., 2009) and metachromatic leukodystrophy (MLD) (Biffi et al., 2006). However, the LV-based gene therapy protocol

developed by Miccio et al was able to correct the β-thalassaemia phenotype with a low vector dosage (less than one provirus copy per cell) in a mouse model (Miccio et al., 2008); providing gene transfer efficacy at clinical vector doses and increasing the biosafety of the protocols.

Genotoxicity (also see section 10). A wide number of preclinical studies have demonstrated the correction of genetic diseases by LV vector-mediated expression of the affected gene (Zhao et al., 2009); (Biffi and Naldini, 2007); (Charrier et al., 2005). Integration of the provirus enables stable transgene expression and propagation down the cell progeny, but it also carries an intrinsic risk of insertional mutagenesis caused by the deregulation of surrounding genes (Montini et al., 2009), (Modlich et al., 2009). However, SIN-LV vector design allows the incorporation of weak cellular or lineage-specific promoters to drive a safer expression of the transgene, thereby reducing the likelihood of trans-activating neighbouring genes (Zychlinski et al., 2008).

Selective advantage. If genetically corrected cells have a selective proliferative advantage, gene therapy can be successful even when only a small number of the transplanted HSCs have been transduced by the vector, like in the case of ADA-SCID (Fischer et al., 2010), SCID-X1 (Hacein-Bey-Abina et al., 2010) and WAS (Dupre et al., 2006), and may even occur without host conditioning (Aiuti et al., 2009). However, this has not been reported for metabolic diseases like chronic granulomatous disease (CGD) (Grez et al., 2011) and diseases affecting the RBCs, such as β-thalassaemia (Miccio et al., 2008), PKD (Meza et al., 2009) and erythroid protoporphyria (EPP) (Fontanellas et al., 2001). This implies that efficient transduction of HSCs is required for the gene therapy of erythroid diseases, which represents an additional challenge. In these cases, strategies based on the selection of corrected cells carrying bicistronic vectors with the therapeutic transgene and a selection gene, like the enhanced green fluorescent protein (EGFP) (Fontanellas et al., 2000; Fontanellas et al., 2001), a drug-resistant gene (Richard et al., 2003); (Trobridge et al., 2012) or CIDs (Richard et al., 2004), have improved the therapeutic effect in mouse models of these diseases.

Appropriate gene expression. Efficient gene therapy protocols must provide appropriate levels of transgene expression, especially when there is no selective advantage of corrected cells, as mentioned above. LV vectors provide more robust expression than γ-RV (Ellis, 2005)

and their design allows internal promoter selection, controlling the expression needed by choosing appropriately an stronger or weaker promoter. For therapeutic applications, physiological promoters, such as those from the human phosphoglyceratekinase (PGK) or elongation factor 1α (EF1α) genes are preferred, and they have been proven to be weaker trans-activators than viral promoters (Zychlinski et al., 2008). In addition, supraphysiological levels led by viral promoters can be cytotoxic.

Transcriptional silencing. Even when efficient gene transfer has been achieved, transgene expression can be silenced due to epigenetic effects, as it has been reported in animal models (Oertel et al., 2003) and in a clinical trial for X-linked CGD (Stein et al., 2010). However, LV vectors have been seen to be more resistant to silencing than γ-RV, probably because their genome contains fewer CpG dinucleotides (Pfeifer et al., 2002), as CpG dinucleotide methylation contributes to vector silencing by histone deacetylation and chromatin condensation. Different transcriptional regulatory elements are currently being used to avoid transgene silencing (Antoniou et al., 2013), including elements with a border or boundary function, such as insulators and scaffold/matrix attachment regions (S/MARs) (Ramezani et al., 2003), elements with dominant chromatin remodelling, like the locus control regions (LCR) (Ramezani and Hawley, 2010), and elements with transcriptional activating functions, such as the ubiquitous chromatin opening elements (UCOEs) (Antoniou et al., 2003).

Immunogenicity of transgenes/vectors. Another concern in autologous HSC gene therapy is the potential recipient immune response against the protein encoded by the transgene that will limit the hematopoietic engraftment (Alonso-Ferreiro et al., 2011). In the case of monogenic diseases, patients carrying null mutations may develop an immune response against the functional transgene product, removing the corrected cells from circulation. However, this problem is overcome in patients expressing truncated proteins or functionally impaired proteins. An anti-vector immune response is also possible and lentiviral vectors are relatively immunogenic, especially when they are expressed in antigen-presenting cells (Brown et al., 2007c). Although this is an advantage in cancer immunotherapy or vaccination, is one of the major drawbacks in gene therapy for metabolic diseases (Escors and Breckpot, 2010). Recently several approaches are focussed on inducing antigen-specific tolerance by *ex vivo* or *in vivo* induced T regulatory cells (Roncarolo and Battaglia, 2007), or on the specific inhibition of transgene expression in antigen presenting cells by including the

mir143-3p target microRNA in the viral vector construction (Brown et al., 2007b), which avoids exogenous protein antigens presentation by the recipient dendritic cells.

8.2 Genotoxicity

The first human gene therapy trials providing therapeutic benefits were performed ten years ago for two types of SCID: ADA-SCID (Aiuti et al., 2002) and X1-SCID (Gaspar et al., 2004). However, the occurrence of T-leukaemias caused by insertional mutagenesis in some X1-SCID patients (Hacein-Bey-Abina et al., 2003); (Howe et al., 2008), highlighted the need to improve vector design and safety. Both trials used early-generation γ-RV based on the Moloney mouse leukaemia virus (MLV), in which transgene expression was driven by the powerful viral LTRs. LTR promoter sequences yield a high efficiency of transduction, but they can affect the regulation of surrounding genes. A high frequency of leukaemia development was observed in transplanted patients from the X1-SCID French trial, due to γ-RV being integrated in close proximity to several proto-oncogene promoters (LMO2 and BMI1), activating some important regulatory genes (such as NOTCH1) or inhibiting tumor suppressor genes (such as CDKN2A) (Hacein-Bey-Abina et al., 2008). In addition, gene therapy attempts for X-linked CGD (X-CGD) based on γ-RV vectors also showed problems associated with vector integration. Two of the X-CGD patients treated in the first clinical trial developed myelodysplastic syndrome (Stein et al., 2010) caused by the insertional activation of MDS1-EVI1 locus (Ott et al., 2006). Further genome mapping studies of γ-RV and LV insertion pattern demonstrated that γ-RV integration preferentially occur near the transcription start sites and promoter regions (Wu et al., 2003), while HIV-1 LV integration occurs along transcriptionally active genes (Schroder et al., 2002); (Mitchell et al., 2004). The fact that LV are less prone to insertional mutagenesis than γ-RV (Trono, 2003), together with the inactivation of the LTR in SIN-LV vectors and their ability to transduce non-diving cells made them replace the early γ-RV in the gene therapy translational field.

Lentiviral vectors have been tested in animal models of a number of hematopoietic diseases, including beta-thalassaemia (Zhao et al., 2009); (Miccio et al., 2008), X1-SCID (Throm et al., 2009), WAS (Mantovani et al., 2009); (Charrier et al., 2005), (Frecha et al., 2008), haemophilia (Johnston et al., 2012); (Brown et al., 2007a), PPE (Richard et al., 2001), congenital erythropoietic porphyria (CEP) (Robert-Richard et al., 2008); metachromatic

leukodystrophy (Visigalli et al., 2010), and Fanconi anemia (Rio et al., 2008). Currently, there are several ongoing clinical trials for the treatment of WAS (Charrier et al., 2007); (Galy et al., 2008), beta-thalassaemia (Cavazzana-Calvo et al., 2010); (Bank et al., 2005), Fanconi anemia (Jacome et al., 2009) and storage disease such as ALD (Cartier et al., 2012; Cartier et al., 2009).

9. Retroviral vectors

Lentiviruses (LV) and gammaretroviruses (γ-RV) are members of the *Retroviridae* viral family, both with a single-stranded (ss) RNA genome and two copies of positive strand RNA per viral particle. Lentiviral vectors have some features that make them suitable for use as therapeutic vectors: 1) their life cycle requires the reverse transcription and integration of their genome into the host genome, providing stable and long-term expression of the viral genes in the infected cells and their progeny (Coffin, 1992). 2) They are able to infect both dividing and non-dividing cells. 3) Lentivirus genomic structure has been widely characterized, enabling its modification. 4) Viral structural proteins can be provided in *trans*, ensuring the ectopic expression only of the transgenes, and increasing the safety of the protocols (Dull et al., 1998) because there is no expression of viral protein after transduction. 5) Envelope proteins can be modified, developing chimeric envelopes and redirecting the tropism of the vector.

Foamy vectors derived from spumaviruses (HFV) have been also used in gene therapy approaches (e.g dog PKD model) (Trobridge et al., 2012). They are apathogenic in humans and able to complete reverse transcription before entering the nucleus, with relatively low preference for integration into active gene sites. However, HFV yield the lowest viral titer and they are the least flexible for pseudotyping among the three different viruses mentioned. Adeno-associated virus 2 (AAV2) has been successfully used in gene therapy for β-thalassaemia (Tan et al., 2001), and they are relatively easy to produce, but the insert size is limited to just 5kb and most of the AAV2 genome remains episomal in the cell.

9.1 General characteristics of lentiviruses

Lentiviruses include primate retroviruses, like the human immunodeficiency virus (HIV; ≈ 9 Kb genome), and most lentiviral vectors validated in pre-clinical studies or already used in clinical trials are based HIV-1. Their genome is organized into three main structural genes: *Gag* (encoding viral matrix, capsid and nucleocapsid proteins that protect the viral genome), *Pol* (encoding a protease, reverse transcriptase and integrase, required for viral replication) and *Env* (encoding the viral surface and transmembrane proteins) (Figure 6.A). In addition to these major proteins shared with γ-RV, complex lentiviral genomes also include the regulatory genes *Tat*, which activates viral transcription, and *Rev,* involved in the control of splicing and nuclear export of viral transcripts (Sakuma et al., 2012). Some of them also have accessory proteins, such as *Vif, Vpr, Vpu* and *Nef*, whose combined action regulates viral gene expression, assembly and replication (Coil and Miller, 2004).

Among all additional proteins, only *Tat* and *Rev* are essential for the expression of viral genes. *Tat* is an activator of the viral promoter necessary for optimal viral RNA production, while *Rev* interacts with the viral RRE element (Rev-responsive element) facilitating the nuclear import of the viral RNA. The viral genome is flanked by long terminal repeats sequences (LTRs) that regulate viral transcription, reverse transcription and polyadenylation of viral mRNA. The 5'-LTR acts as a combined enhancer and promoter for transcription by host RNA polymerase II, while 3'-LTR stabilizes these transcripts by mediating their polyadenylation (Delenda, 2004). The packaging sequence (Ψ) is located between the 5'-LTR and the *Gag* gene (Figure 7) and is required for packaging the viral genome into the viral capsid.

The retroviral life cycle (Figure 6) starts with the binding of the viral envelope to its receptor in the target cell, and the subsequent fusion of the viral envelope with the membrane (step 1 Figure 6). At that moment, the viral particle is internalized and the viral core is released into the cytoplasm (step 2). The ssRNA is reverse transcribed into dsDNA within the core (3), which is transported to the nucleus (step 4) upon cell division for gammaretrovirus (simple retrovirus) or through active transport in the case of lentiviruses (Escors and Breckpot, 2010). Once in the nucleus, the viral DNA is integrated into the host genome (provirus) (step 5) yielding long-term expression of viral genes, which are transcribed (step 6) and spliced (step 7) during the life of infected cells. The full-length viral

RNA and the RNA coding for the viral structural proteins are transported to the cytoplasm (step 8) to start the translation (step 9). The unspliced viral RNA is packaged and the viral particle is assembled (step 10), which matures and escapes from the cell (step 11) becoming an infectious viral particle.

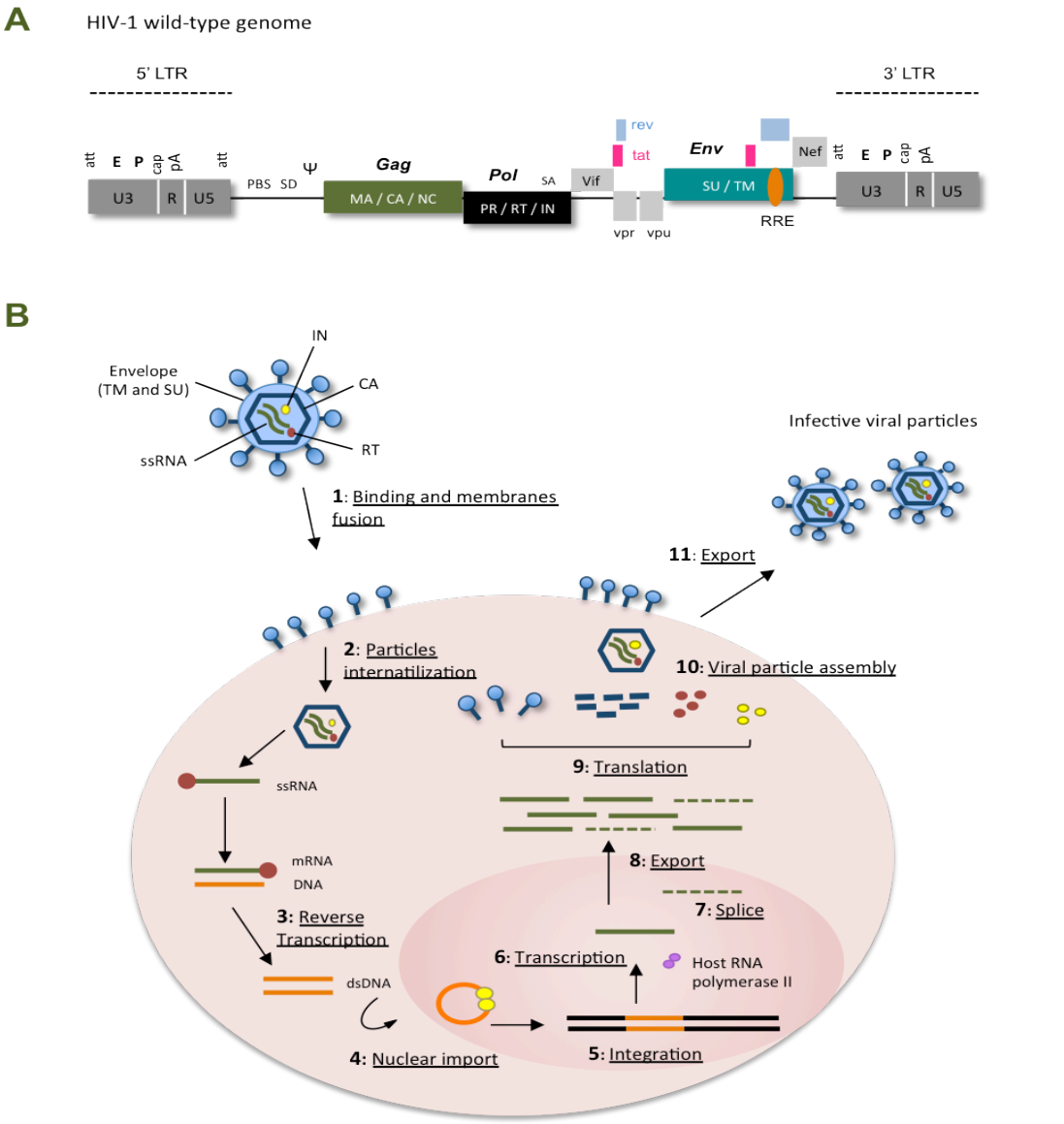

Figure 6: Retroviral genomic structure and life cycle. **A)** HIV-1 lentivirus wild-type genome, showing the *Gag, Pol* and *Env* structural genes, the regulatory genes *Tat* and *Rev*, and the accessory genes *Vif, Vpr, Vpu, Nef*). LTR: long-terminal repeats, E: enhancer, P: promoter, att: integration sites, cap: 5'-5' joining between the first 5' nucleotide of the RNA and a methylated ribonucleotide, pA: polyadenilation signal, PBS: primer binding site (binding site for the tRNA to start the reverse transcription), SD: splice donor, SA: splice acceptor, Ψ: packaging signal, MA: matrix, NC: nucleocapsid, PR: protease. **B)** Retroviral life cycle. IN: integrase, CA: capsid, RT: reverse transcriptase, TM: transmembrane protein, SU: surface protein. Modified from (Baum et al., 2006).

9.2 Production of self-inactivating lentiviral vectors

Currently, HIV-1-based lentiviral vectors are widely used in the gene therapy field (62 clinical trials, 3.3% of the total ongoing clinical trials; http://www.wiley.co.uk/genmed/clinical/ [March 2013]), and they have replaced the early γ-RV vectors used in the first experiments because of their capability of transducing quiescent cells and their safer characteristics (Modlich et al., 2009). Moreover, LV can carry larger transgenes (up to 10 kb) than γ-RV, athough vector titers tend to decrease when larger transgenes are used (Sinn et al., 2005). To minimize the risk of generating replication-competent lentiviruses (RCLs), LV production methods have been focused on reducing the homologous sequences between vector and the helper sequences (structural or regulatory genes) by splitting the helper functions into multiple plasmids. Thereby, final viral particles will only contain the transcripts with all *cis*-regulatory sequences required for its packaging Figure 7, but not the *trans*-acting genes that are carried by the helper plasmid. In general, LV particles are generated by co-transfection of 3 plasmids in human embryonic kidney (HEK) 293T cells (Naldini et al., 1996): the packaging plasmid (*Gag*, *Pol*, *Tat* and *Rev* proteins), the transfer plasmid (transgene expression cassette) and an envelope-encoding plasmid (Figure 7).

The minimal transgene expression cassette contains all the *cis*-active sequences needed for packaging (Ψ), reverse transcription (LTRs and primer binding site, PBS), integration (attL and attR integration sites) and transcription (5'LTR or internal promoters), as well as the transgene or therapeutic gene. However, as HIV-1 is a human pathogen, other modifications in LV production methods have been performed in order to prevent possible RCLs generation and to improve viral titer and transgene expression (Delenda, 2004). Some of them are described below:

Packaging plasmid: accessory viral genes (*vif, vpr, vpu* and *nef*) have been removed from the packaging plasmid, reducing the potential recombination with wild-type viruses (Zufferey et al., 1997). The latest LV vectors split tat-rev genes into two plasmids to reduce the possibility of homologous recombination, as only three (*Gag-Pol-Rev*) of the nine viral genes direct the packaging of new viral particles (Dull et al., 1998).

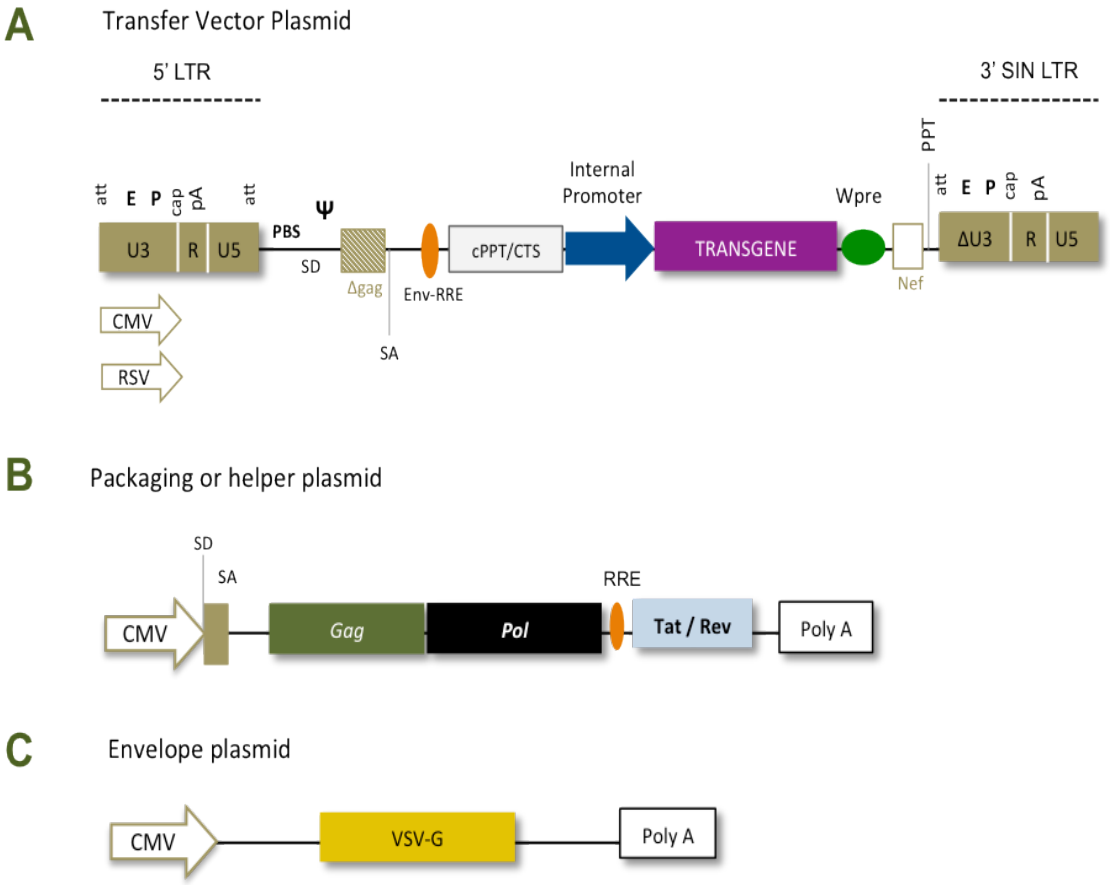

Figure 7: Split-packaging system for LV vector production. **A)** Lentiviral design with the necessary *cis*-genes and other additional elements to improve viral titers and safety. The substitution of the 5'LTR U3 region by CMV (pCCL backbone) or RSV (pRRL backbone) promoters allows viral production independent of *Tat*. **B)** Packaging and **C)** envelope plasmids provide in *trans* the rest of the viral proteins required for LV vector production. Modified from (Baum et al., 2006) and (Sinn et al., 2005).

Transfer vector: both LTR sequences have been modified to generate self-inactivating (SIN) LV vectors.

- SIN vectors have a 400 bp deletion in the 3'-LTR, covering the promoter/enhancer elements from the U3 region. Expression of the transgene is thereby dependent on internal promoters, reducing the risk of RCLs and decreasing promoter interference (Ginn et al., 2003). This 3'-LTR deletion removes the TATA box, preventing transcription initiation (Miyoshi et al., 1998); (Zufferey et al., 1998) and therefore inactivating the vector.
- The U3 region of the 5'-LTR has been replaced by other heterologous promoting sequences (i.e. CMV or RSV) to achieve a Tat-independent

transcription and to increase genomic RNA synthesis, resulting in the increase of the viral titer. Because 5'-U3 region drives the expression of primary transcripts, its modifications will not be present in transduced cells (Schambach et al., 2009) (Figure 8).

- Exogenous elements, such as β-globin or SV40 polyadenylation signals (Iwakuma et al., 1999) or the upstream sequence element (USE) from simian virus 40 (SV40-USE) (Schambach et al., 2007), have also been included in the R region of the viral 3'LTR in order to decrease the transcriptional readthrough from the internal promoters (Zaiss et al., 2002) or from remnants of the deleted U3 region of SIN-LV vectors (Almarza et al., 2011), preventing the potential transcriptional activation of the downstream genes.
- The leader region contains the packaging signal (Ψ), and LV vectors were thought to require approximately 300 bp of the *Gag* gene in this region, although this has now been reduced to just 40 bp.
- Rev responsive element (RRE) has also been included to improve the efficiency of gene transfer, although it contains surrounding *Env* remnants, which may bring safety concerns.
- The central polypurine tract (cPPT), which facilitates nuclear translocation of the pre-integration complexes, together with the central terminal sequence (CTS) involved in the separation of reverse transcriptase, has been seen to improve viral titers (Zennou et al., 2000); (Follenzi et al., 2000).
- The woodchuck hepatitis virus (WHV) post-transcriptional regulatory element (Wpre) significantly increases transgene expression in target cells, by increasing RNA stability in a transgene, promoter and vector-independent manner (Zufferey et al., 1999). However, it can express a truncated 60-amino acid protein derived from the WHV X gene involved in liver cancer (Kingsman et al., 2005). Therefore, most pre-clinical protocols and clinical trials include a mutated version of the Wpre element (Zanta-Boussif et al., 2009). On the other hand, the use of two SV40-USE elements in SIN-LV vectors has been seen to be more efficient than Wpre sequence in supressing transcriptional readthrough (Schambach et al., 2007).
- dNEF/PPT signal is essential for reverse transcription, and its incorporation significantly improves LV vector production.

- The use of codon optimized transgenes helps to achieve higher expression levels. To prevent unpredictable splicing, the transgene usually contains only the coding sequence without introns, and the *in silico* splice site prediction avoids cryptic splice sites, which could result in the loss of important LV vector sequences (Schambach et al., 2009). In addition, optimized sequences have a high GC content to increase mRNA stability (see section 11.1.B of Materials and Methods).

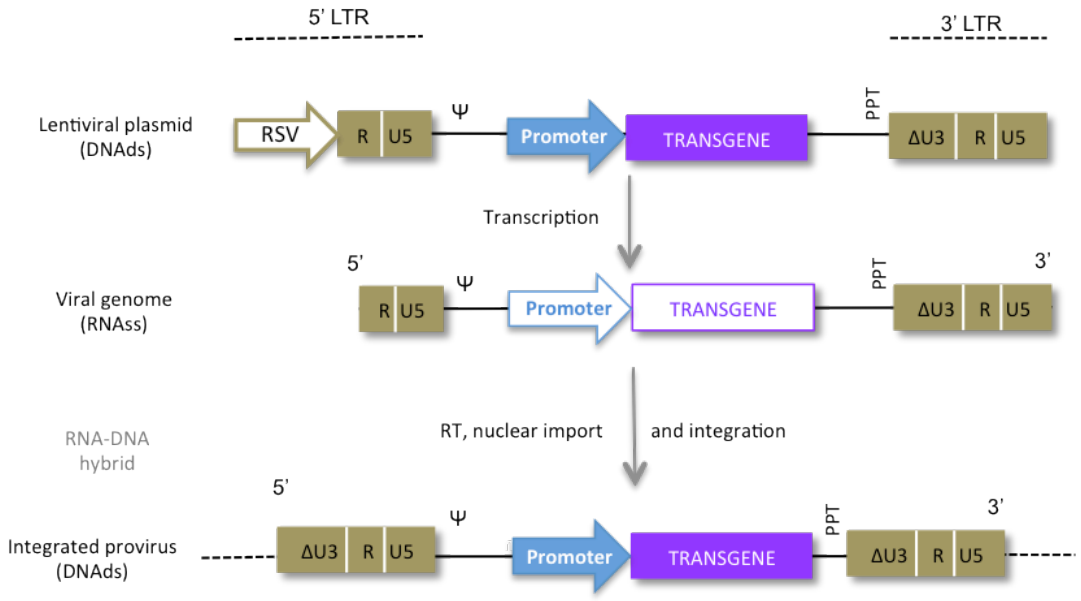

Figure 8: Schematic representation of transcription and integration of the LV vectors. Two main characteristics of the self-inactivating LV design allow the modulation of levels and specificity of transgene expression, reducing at the same time the likelihood of transactivating surrounding genes. First, the substitution of the 5'LTR U3 region in the transfer plasmid by CMV or RSV promoters improves the viral titer, as these viral promoters will drive the expression of the primary transcripts in the packaging cell line. Second, the deletion of the enhancer/promoter sequences of the 3'LTR U3 region allows the control of transgene expression by the internal promoter because upon reverse transcription and integration of the provirus within the target cell genome, the transcription unit will be flanked by the 3'LTR at both 3' and 5' ends.

Despite these safety modifications, the minimal packaging signal of the transfer vector still contains 40 bp of the *Gag* gene, the cPPT tract comprising of 130 bp derived from *Pol*, and the presence of RRE element including *Env* remnants, may still be a safety threat. Because of that, additional modifications have been performed in the packaging system to obtain Rev/RRE-independent LV vectors or reduce aberrant transcription by using cryptic transcriptional activators (Baum et al., 2006).

Envelope plasmid: The most widely used LV vector envelope is the attachment glycoprotein of the vesicular stomatitis virus (VSV-G), which allows the production of high-titer vectors and confers a broad host range (Hanawa et al., 2002) (Follenzi et al., 2000) (Guenechea et al., 2000). Modification of the envelope protein by adapting other viral envelopes (e.g Glycoprotein gp70 from Gibbon ape leukaemia virus, GALV), by truncation of the intracellular domain (e.g GALV-TR), or by the introduction of cell-type specific ligands or antibodies (GALV-TM) (Yang et al., 2006), can redirect the binding of the LV particle to a specific target cell. Although envelope engineering may alter the fusion domain of *Env,* resulting in low viral titers. Downstream processing of lentiviral supernatants (e.g ultracentrifugation) reduces the toxicity of VSV-G and cell debris, although it is still susceptible to serum inactivation. The engineered envelope RDTR/SCFHA-LV (RD114/SCF) has been seen to be more resistant to degradation by human complement and pseudotyped LV vectors efficiently transduce very immature $hCD34^{+}$ HSCs (Frecha et al., 2011).

Transient transfection of packaging cell lines can produce high viral titers but it is difficult to scale up. Stable packaging cell lines have been developed to overcome this limitation, but both lentiviral protease (encoded by the *Pol* gene) and VSV-G envelope are cytotoxic, so other strategies are being tested, such as inducible promoters within a Tet-On Tet-Off system (Matrai et al., 2010).

10. Safety issues of gene therapy

Genetic modification of HSCs using lentiviral vectors has brought encouraging results in the correction of genetic diseases, but the risk of insertional mutagenesis is still present (Modlich et al., 2009). Genotoxicity is one of the major limitations in the gene therapy field, due to potential hematological malignancies derived from the vector integration. Since several X1-SCID patients from the γ-RV clinical trial developed leukaemia (Hacein-Bey-Abina et al., 2003) (Hacein-Bey-Abina et al., 2008), a lot of efforts have been focused on increasing the safety of gene therapy protocols through improving both vector design and methods to detect vector integrations in the cell genome.

Although LV and γ-RV exhibit semi-random integration profiles, SIN-LV vectors have been proven to have lower oncogenic potential than γ-RV vectors (Baum et al., 2007) (Bushman et al., 2005) because while SIN-LV vectors are more likely to integrate into active transcription units, γ-RV vectors have a preference towards integrating near promoters of active genes and around DNase I hypersensitive sites (HS), increasing their genotoxic risk (Mitchell et al., 2004) (Wu et al., 2003) (Schroder et al., 2002). The most dangerous insertions are those leading to gain-of-function mutations, such as activation of a proto-oncogene flanking an insertion site, as well as mutations that activate cell cycle regulatory genes or loss-of-function mutations that inhibit tumor-supressor genes (Hacein-Bey-Abina et al., 2008); (Howe et al., 2008); (Boztug et al., 2010). Also, studies carried out with MLV-derived γ-RV vector and simian immunodeficiency virus (SIV) LV vector showed that HSCs were particularly susceptible to insertional mutagenesis because of the expression of genes involved in self-renewal control and proliferation before HSCs start to differentiate towards a specific hematopoietic lineage (Hematti et al., 2004).

Either enhancers/promoters present in the vector or aberrant splicing from the vector transcripts can mediate insertional mutagenesis (Figure 9). In γ-RV with active LTRs, aberrant transcripts can be generated from the 5'LTR or by transcriptional readthrough from the 3'LTR, and promoter sequences within the 5'LTR can potentially transactivate cellular promoters located close to the integration site (Naldini, 2011). SIN-LV vectors display lower insertional mutagenesis potential because of the 400 bp deletion in the 3'LTR U3 region, although polyA signals within the 3'LTR can still lead to aberrant transcripts, and the internal promoter could transactivate neighbouring cellular promoters (Wu et al., 2003).

Comparative studies were carried out to evaluate the genotoxicity of SIN-LV and γ-RV in both human and mouse HSCs. Montini et al developed an *in vivo* genotoxicity assay by transduction and transplantation of HSCs in the $Cdkn2a^{-/-}$ tumor-prone mouse model (Montini et al., 2006). The analysis of integration site selection and tumor development demonstrated that, unlike SIN-LV vectors, γ-RV integrations accelerate tumor onset in transplanted mice. However, in this study SIN-LV vectors were compared with non-SIN RV vector, and it has been proven that SIN-RV vectors have a much safer profile than their non-SIN counterparts (Cornils et al., 2009). Further studies in this model have confirmed the lower genotoxic potential of LV over γ-RV (Montini et al., 2009), although the sensitivity of this genotoxic assay may be decreased by the genetic instability of this model, as even SIN-

LV with a strong internal promoter, such as the spleen focus-forming virus (SFFV) enhancer-promoter, were not able to trigger transformation above the background levels of this model.

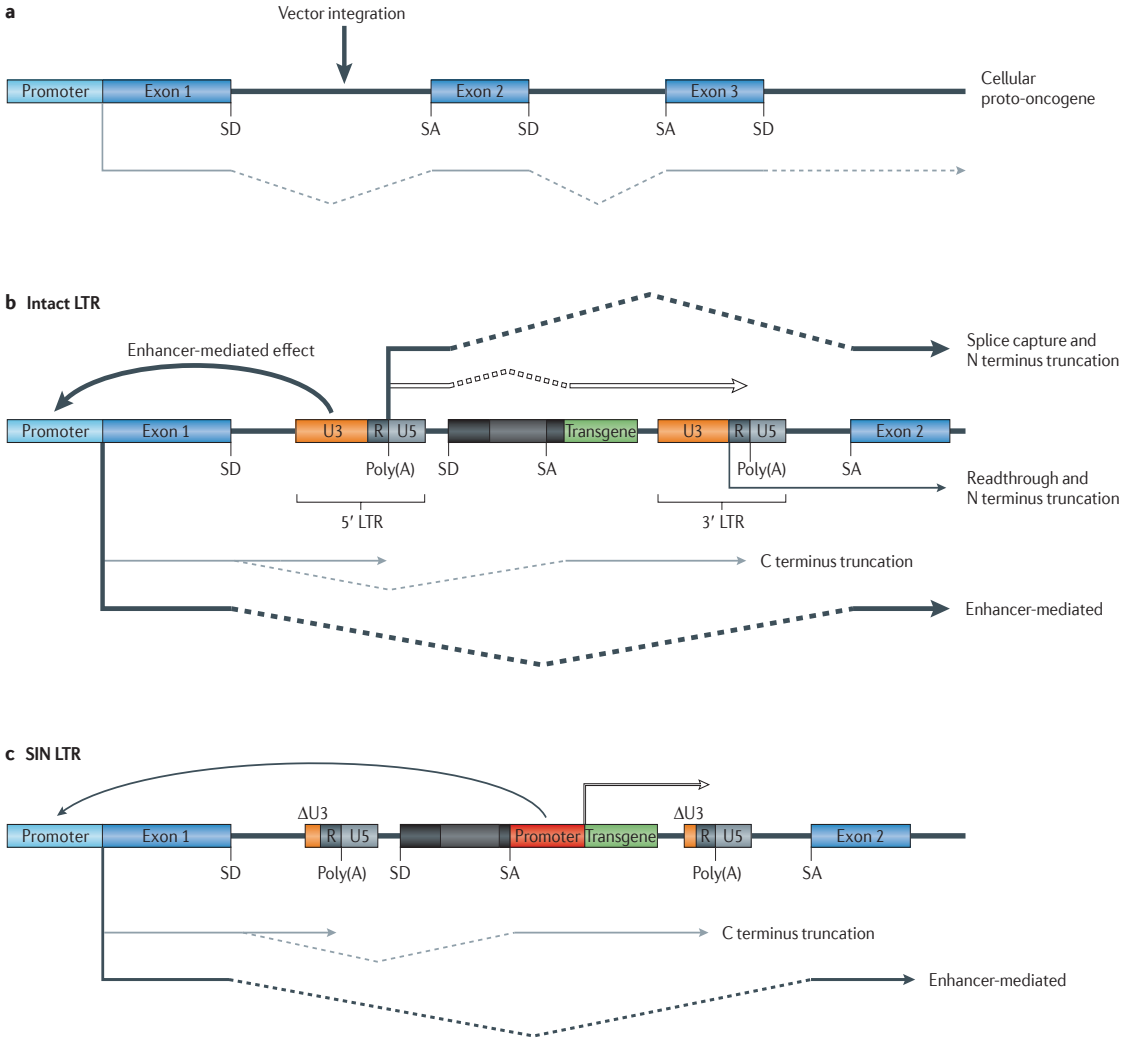

Figure 9: Mechanisms of vector insertional mutagenesis. **A)** A vector's integration site is shown in the first intron of a proto-oncogene. The grey arrow indicates the normal transcript, and the broken segments represent intronic sequences removed by splicing. **B)** A conventional γ-RV is integrated at the site indicated in A). The white arrow represents the vector transcript that encodes the transgene, but cellular promoter, 5'LTR promoter region and inefficient transcription termination can deregulate the proto-oncogene expression by enhancer-mediated effects, alternative splice capture or transcriptional readthrough, respectively. **C)** In SIN-LV vectors all LTR-dependent events are avoided, but a residual risk of long-range transactivation of the oncogene mediated by the enhancer elements of the internal promoter is still present. This can be overcome by the use of promoters with moderate to weak activity (from (Naldini, 2011).

At the same time, Modlich *et al* developed an *in vitro* cell immortalization assay (IVIM) for genotoxicity evaluation, observing that the pattern of LV insertion was approximately threefold less likely than γ-RV to activate transformation genes in primary hematopoietic cells (Modlich et al., 2009). Their results supported the previously reported lower genotoxic potential of LV vectors, but also highlighted the potential insertional mutagenesis risk of SIN-LV vectors. With the IVIM method they observed that a common integration site (CIS) in the first intron of *Evi1* proto-oncogene had been selected in both LV and γ-RV vectors. Nevertheless, the use of internal promoters with weak to moderate activity, such as EF1α and PGK, significantly decreased the risk of insertional transformation (more than tenfold) in a SIN-LV context (Zychlinski et al., 2008), and no activation of the *Evi1* proto-oncogene was detected even with multiple integrations of the SIN-EF1α vector. Moreover, the use of internal tissue- or lineage-specific promoters provides additional benefits, as we describe below.

11. Erythroid-specific vectors

As mentioned before, self-inactivating lentiviral design allows the use of internal promoters, offering a more physiological expression and less genotoxic profile when using weak promoters. Restricted expression of the transgene to the affected cell lineage by using tissue-specific promoters increases the safety of LV vectors, decreasing side effects in non-target cells. Gene therapy for RBC disorders, such as PKD, could benefit from the use of erythroid-specific vectors, avoiding potential cytotoxic effect in hematopoietic cells other than RBCs. Additionally, physiologically regulated vectors carrying only the promoter and the enhancer sequences of the affected gene increase expression stability and reduce gene silencing (Toscano et al., 2011). Most erythroid-specific promoters and enhancer sequences have been tested in hemoglobinopathies, and regulatory sequences from the human beta-globin gene have been widely used to increase erythroid-specificity (Miccio et al., 2008); (Antoniou et al., 2013).

11.1 Erythroid tissue-specific promoters

Different combinations of erythroid-lineage specific promoters and enhancers have been tested in SIN-LV approaches in order to yield efficient and restricted expression in the target cells. Moreau-Gaudry et al compared the efficiency of EGFP transgene expression with four different erythroid promoters: ankyrin-1, α-spectrin, β-globin and ζ-globin, and four different enhancers: GATA-1 autoregulatory element (RE), β-globin Locus Control Regions (LCR), intron I8 from the 5-aminolevulinate synthase (*hALAS*) gene and α-globin HS40, in both human and mouse HSCs, as well as in engrafted animals. They found that the highest levels of expression were obtained when the ankyrin-1 promoter was combined with either GATA-1/HS40 or I8/HS40 enhancers (Moreau-Gaudry et al., 2001). On the other hand, enhanced erythroid expression has been also reported in two mouse models of metabolic erythroid diseases: EPP (Richard et al., 2001) and CEP (Robert-Richard et al., 2008), using SIN-LV vectors that carried the ferrochelatase (FECH) or the uroporphyrinogen III synthase (UROS) cDNA, respectively, and the chimeric promoter ankyrin-1/ α-globin HS40 (Table 3).

11.2 Human β-globin locus control regions

The addition of positive regulatory elements from the human β-globin gene has improved the prospects of gene therapy for RBC diseases, by increasing the erythroid specificity of the vectors. LCRs are tissue-specific regulatory elements that confer physiological levels of expression in a transgene copy number-dependent manner (Li et al., 2002). In humans, β-LCR regulates the synthesis of *HBB* family (β-globin gene), yielding erythroid and developmental-stage specific β-globin expression at high levels. β-LCR consists of five elements (HS1-HS5) of the human ε-globin gene that contain a high degree of DNaseI hypersensitive site (HS sites) (Grosveld, 1999). Each HS site consists of a 200-400 bp core of nucleosome-free regions of open chromatin, highly accessible to *trans*-acting factors (Forrester et al., 1986) because of the high density of binding sites for ubiquitous (Sp1 and USF) and erythroid-specific (GATA-1, EKLF, NF-E2 and Tal1) transcription factors (Liang et al., 2008). Elements of HS1-HS4 posses strong transcription-enhancing activity (Antoniou and Grosveld, 1990), while HS5 is a developmental stage-specific insulator element (Palstra et al., 2008). The mechanism of controlling β-globin gene expression by β-LCR is not completely

understood, but it is thought that β-LCR would act as an open chromatin structure making the DNA more accessible for transcription factors (Liang et al., 2008).

Its use to increase transgene expression in the erythroid lineage has been proven in a number of preclinical studies based on SIN-LV vectors, mainly β-thalassaemia *intermedia*, β-thalassaemia major and sickle cell disease (SCD), with the correction of the disease phenotype (Table 3). Some of the strategies described in Table 3 for achieving erythroid-specific expression of globin genes could be adapted for vectors directed to other erythroid diseases, such as PKD. In addition, although *in vitro* and *in vivo* studies have shown that β-LCR can enhance erythroid-specific expression from heterologous non-erythroid promoters (Blom van Assendelft et al., 1989) (Montiel-Equihua et al., 2012), it is important to use both β-LCR enhancers and erythroid-promoters for obtaining erythroid-specificity in gene therapy protocols for RBC diseases. This approach has been successfully performed in a mouse model of haemophilia, achieving high levels of clotting factor IX in the circulation, produced by RBCs after transduction of HSCs with lineage-restricted LV vectors (Sadelain et al., 2009).

The transcriptional activating potential of β-LCR, even at long distances, raises the possibility of insertional mutagenesis following proviral integration. Indeed, β-LCR-based LV vectors have been shown to activate host genes in erythroid cells *in vitro* (Arumugam et al., 2009a) and *in vivo* (Hargrove et al., 2008) at up to 100 kb from the site of the integration site, but the frequency was very low and none of the affected genes were associated with cancer development (Arumugam et al., 2009a). In addition, β-LCR enhances transcription from the nearest promoter susceptible to be activated, in this case the internal promoter of the LV vectors, likely displaying a very low propensity to activate host genes (Antoniou et al., 2013).

11.3 Insulators

An additional improvement to provide safer vectors for RBC gene therapy is the use of insulator elements, which have been seen to protect expression cassettes from positional effects, and therefore decrease the risk of transactivating neighbouring endogenous genes. Insulators are genomic elements that can shelter genes from their surrounding chromosomal environment by either blocking the action of a distal enhancer on the

Erythroid tissue-specific SIN-LV vectors

Disease	Enhancer	promoter	transgene	Other sequences	mouse model	Reference
–	GATA-1 RE β-LCR I8 *hALAS* HS40	β-Globin ζ-Globin Ankirin-1 α-spectrin	EGFP	Wpre	*in vitro* and engrafted C57BL/6	Moreau-Gaudry et al., 2001
EPP	HS40	Ankirin-1	*FECH* cDNA	Wpre	*Fech*$^{m1pas/m1pas}$	Richard et al., 2001
CEP	HS40	Ankirin-1	*UROS* cDNA	Wpre	*Uros*$^{mut248/mut248}$	Robert-Richard et al., 2008

Physiologically regulated erythroid SIN-LV vectors

β-thalassaemia intermedia	3.2 Kb β-LCR (HS2, HS3, HS4)	hβ-Globin	hβ-Globin (*HBB* cDNA)	β-globin 3' enh	*Hbb*$^{Th3/+}$	May et al., 2000 May et al., 2002
	≈ 2 Kb β-LCR (HS2, HS3, HS4)	hβ-Globin	hγ-Globin (*HBG* cDNA)	β-globin 3' enh or γ-Globin 3' RE	*Hbb*$^{Th3/+}$	Persosn et al., 2003
	3.2 Kb β-LCR (HS2, HS3, HS4)	hβ-Globin	hγ-Globin (*HBG* cDNA)	γ-Globin 3'UTR	*Hbb*$^{Th3/+}$	Hanawa et al., 2004
	2.7 Kb β-LCR [A] (HS2 and HS3)	hβ-Globin	hβ-Globin (*HBB* cDNA)		*Hbb*$^{Th3/+}$	Miccio et al., 2008
	3.2 Kb β-LCR[B] (HS2, HS3, HS4)	hβ-Globin	hβ-Globin (*HBB* cDNA)	Wpre Ank ins in 3'ΔLTR	*Hbb*$^{Th3/+}$	Breda et al., 2012
β-thalassaemia major	3.2 Kb β-LCR (HS2, HS3, HS4)	hβ-Globin	hβ-Globin (*HBB* cDNA)	β-globin 3' enh	*Hbb*$^{Th3/Th3}$	Rivella et al., 2003
	3.1 Kb β-LCR[C] (HS2, HS3, HS4)	hβ-Globin	hβ-Globin (*HBB* cDNA)	β-globin 3' enh cHS4 ins in 3'ΔLTR	Humanized NOD-SCID	Puthenveetil et al., 2004
SCD	2.6 Kb β-LCR (HS2, HS3, HS4)	hβ-Globin	hβ-Globin (*HBB* cDNA)	β-globin 3' enh	BERK and SAD	Pawliuk et al., 2001
	HS2, HS3, HS4	hβ-Globin	hβ/γ-Globin (hybrid gene)		BERK	Perumbeti et al., 2009
	3.1 Kb β-LCR HS2, HS3, HS4	hβ-Globin	hγ-Globin (*HBG* cDNA)	β-Globin 3'UTR or γ-Globin 3'UTR	BERK	Pestina et al., 2009

A) This vector efficiently transduced β-thalassaemic patient CD34+ cells at high frequency (Roselli et al., 2010) restoring the HbA synthesis.

B) This vector efficiently transduced β-thalassaemic and SCD patient CD34+ cells, producing potential therapeutic levels of HbA (Breda et al., 2012)

C) This vector provided normal levels of β-globin expression in erythroid cells derived from β-thalassaemic patient BM (Malik et al., 2005)

Table 3: Specific self-inactivating lentiviral vectors for gene therapy of erythroid diseases. Table summarizes some of the elements used in SIN-LV designs to achieve erythroid-specific expression of different transgenes. These vectors have been successfully used to correct several RBC genetic diseases, in both patients CD34+ cells and mouse models.

promoter, or by acting as barriers that protect the gene from the silencing effect of heterochromatin (Lisowski and Sadelain, 2008). The most studied element is the chicken hypersensitive site 4 (cHS4) from the β-globin gene (Arumugam et al., 2007), which has been included in the 3'LTR of the SIN-LV vectors, together with the βLCR in a preclinical study for β-thalassaemia (Table 3). In this protocol, the authors used the full genomic 1.2-kb cHS4, which severely inhibited LV production (Puthenveetil et al., 2004), but a reduced sized 650-bp version of cHS4 element has been successfully tested (Arumugam et al., 2009b), being especially functional in erythroid cells. In addition, cHS4 insulator has been proven to confer a sustainable expression in human HSCs (Ramezani et al., 2003). Recently, other insulators are being studied, such as the 190 bp ankyrin insulator, which efficiently corrected the β-thalassaemia phenotype in $Hbb^{Th3/+}$ mice and increased the β-globin expression in $CD34^{+}$ cells from β-thalassaemia and SCD patients (Breda et al., 2012).

The last evidence of the usefulness of β-LCR sequences comes from the gene therapy clinical trial for β-thalassaemia led by Philippe Leboulch (Cavazzana-Calvo et al., 2010), based on a SIN-LV carrying the HS2, HS3 and HS4 sites and the cHS4 insulator. Through transduction and transplantation of corrected HSCs, one β-thalassaemic patient has become transfusion independent. However, this outcome represents only a partial success of gene therapy for an erythroid disease such as thalassaemia, because most of the therapeutic benefits result from a dominant, myeloid cell clone in which the integrated vector causes transcriptional activation of *HMGA2* gene in erythroid cells. This gene is expressed as a truncated mRNA insensitive to degradation by let-7 microRNA, leading to therapeutic efficacy (Cavazzana-Calvo et al., 2010).

12. GENE THERAPY FOR PKD

The proof-of-principle that a gene repair strategy could be a feasible option to treat PKD disease was first provided by Tani *et al*. They introduced the human LPK cDNA into C57BL/6 mouse BM cells using a γ-RV vector carrying the SV40 promoter. Thirty days after transplant, hLPK was expressed in the peripheral blood of C57BL/6 recipients (Tani et al., 1994), but long-term expression of the enzyme was not detected in the hematopoietic compartments. Later, a transgenesis gene addition approach to correct the PKD phenotype

was addressed using the CBA-*Pk-1*^{slc}/*Pk-1*^{slc} deficient mouse model, based on the pronuclear injection of a DNA construction carrying the h*PKLR* gene. In this assay, Kanno et al developed a minigene containing a genomic fragment covering exons 1 and 2 of the human *PKLR* gene (proximal promoters of both RPK and LPK), and the fragment spanning from exon 3 to the 3'-untranslated region (3'UTR). They also included a 3.1 kb fragment of the human β-LCR regulatory sequences. Despite the authors demonstrating the full correction of the hemolytic anemia and reticulocytosis in homozygous mice expressing high levels of hRPK via transgenesis, this strategy did not revert splenomegaly (Kanno et al., 2007). Notably, they found a decreased apoptosis in spleen erythroid compartment of corrected mice, suggesting that hRPK ectopic expression is able to rescue deficient erythroid progenitors from apoptosis. A recent work has reported therapeutic benefits in a PKD Basenji dog by *in vivo* expansion of corrected hematopoietic repopulating cells. This approach was based on a tri-cistronic foamy vector that expressed EGFP, the canine RPK and the P140K mutant of the methylguanine methyltransferase (MGMTP140K) drug resistant gene (Trobridge et al., 2012). With this strategy, Trobridge *et al* have achieved an increase in the percentage of corrected cells (from 3.5% to 33%) resulting in a reversion of the canine PKD disease with stable hematocrit, stable reduction of reticulocytes, lower lactate dehydrogenase (LDH) levels and transfusion independence.

Our previous work carried out in the AcB55 mouse model also supported the therapeutic potential of viral vectors for the gene therapy of PKD. We first developed a bicistronic γ-RV vector carrying the hRPK cDNA and the EGFP reporter gene split by a eukaryotic internal ribosome entry site (IRES) element. Through the transduction of HSCs and transplantation into wild-type mice, we achieved high hRPK activity in RBCs and high proportions of erythroid precursors and mature erythrocytes in BM and spleen (Meza et al., 2007). Later, we reported a successful gene therapy approach using the same γ-RV vector in the congenital RPK-deficient mouse strain AcB55 deficient in RPK. Human RPK retroviral-derived expression was capable of fully resolving the pathological PKD phenotype in terms of hematological parameters, reticulocytosis and splenomegaly, together with the normalization of the BM and spleen erythroid progenitors, Epo levels, PK activity and ATP levels. In addition, despite the strong viral LTR to drive the transgene expression, no deleterious effects were observed in secondary and tertiary recipients (Meza et al., 2009). We also observed that more than 25% of genetically corrected cells were needed to fully

revert the RPK deficiency, suggesting that gene therapy protocols for treating PKD will always require a significant number of donor gene-modified HSCs.

As mentioned above, previous works have demonstrated the feasibility of gene therapy for PKD, but efficient transduction and donor chimerism are still the major limitations. In addition, the risk of insertional mutagenesis should be reduced by the use of new vector designs, and the obvious benefits of tissue-specific transgene expression imply that the ideal vector for PKD gene therapy would be a self-inactivating lentiviral vector harbouring an erythroid-specific promoter or at least a weak promoter driving the expression of the hRPK gene.

IV. OBJECTIVES

Pyruvate kinase deficiency is an autosomal recessive disorder caused by mutations in the *PKLR* gene that affect the erythroid metabolism producing chronic nonspherocytic hemolytic anemia (CNSHA). Reduced levels of RPK (R-type PK) lead to an accumulation of glycolytic intermediates that ultimately shortens the life span of mature RBCs. Pathological manifestations usually appear when enzyme activity falls below 25% of the normal level of PK activity. Clinical symptoms of PKD vary considerably from mild to severe anemia, and the disease can be fatal in early childhood, even leading to *in utero* death of very severely affected patients. At present, PKD treatment is based on supportive therapy, such as periodic blood transfusions and splenectomy, often impairing the patient's quality of life.

Currently, the only possible curative treatment is allogeneic bone marrow transplant, and indeed it was successfully performed in one severely affected child. However, the availability of HLA-compatible donors limits the number of patients that can benefit from this treatment, and bone marrow transplant-derived complications can be as severe as the enzymatic deficiency. Giving its recessive inheritance trait and the restriction of the enzymatic defect to erythroid cells, PKD could be a suitable disease to be treated by autologous transplantation of gene therapy corrected HSCs. This strategy could be a therapeutic option for severe cases of PKD applicable, in theory, to all patients, providing a supply of functional blood cells able to rescue the PKD pathology.

In order to have a better understanding of the pathological mechanism involved in PKD disease, and to develop a new treatment based on the genetic modification of HSCs, we considered the following aims:

1. To study the phenotypic characteristic of a *Pklr* deficient mouse strain (AcB55) in order to determine its usefulness as an experimental model of the human PKD disease.

2. To develop a clinically applicable lentiviral vector with ubiquitous expression of the human wild type *PKLR* therapeutic transgene to perform preclinical gene therapy assays, evaluating the efficacy and safety of the vector.

3. To develop erythroid tissue-specific lentiviral vectors to provide new RPK expression-restricted vectors for future gene therapy protocols focussed on PKD.

V. MATERIALS AND METHODS

1. Experimental animals

The mouse strain used in this work was the recombinant congenic mouse AcB55 (also recorded as the PKD mouse), which was obtained from Emerillon Therapeutics Inc. (Montreal, Quebec, Canada) and maintained at the Laboratory Animal Facility of CIEMAT (registration number ES280790000183) by inbreeding. AcB55 strain was originally obtained by systematic inbreeding of two independent strains: C57BL/6 (B6) as donor parental strain and A/J (A) as background parental strain (Mullerova, 2004), therefore carrying 12.5% of the donor genome and 87.5% of the background genome fixed on chimeric chromosomes.

AcB55 recombinant congenic mice carry the malaria susceptibility alleles at *Char1* and *Char2* loci from A/J, and the resistant B6-derived locus *Char4* on chromosome 3 (Fortin et al., 2001), where a point mutation in *pklr* gene was identified. AcB55 mice are homozygous for this loss-of-function mutation that causes the 269T>A substitution in the *Pklr* gene. The consequent isoleucine to asparagine substitution at residue 90 of the PKLR protein (I90N) (Min-Oo et al., 2003) causes hemolytic anemia, acute splenomegaly and a high number of circulating reticulocytes (Min-Oo et al., 2004), resulting in a mouse model of PKD. We also generated hybrid mice derived from C57BL/6 and A/J parents obtained from The Jackson Laboratory. The hybrid (A/JxC57BL/6) F1 or AB6F1 mice were used as healthy controls in the phenotype characterization studies, together with C57BL/6 and A/J mice.

Animals were allowed food and water *ad libitum*, and were routinely screened for pathogens in accordance with Federation of European Laboratory Animal Science Associations (FELASA, Tomworth, United Kingdom) procedures, with no pathogens found. All experimental procedures were carried out according to European and Spanish laws and regulations (European convention ETS 1 2 3, about the use and protection of vertebrate mammals used in experimentation and other scientific purposes and Spanish Law 32/2007, R.D. 1201/2005 and RD 53/2013 about the protection and use of animals in scientific research). Procedures were approved by our Animal Experimentation Ethical Committee according to all external and internal biosafety and bioethic guidelines. Procedures involving Genetically Modified Organism were approved by Spanish Biohazard Commission (Registration Number A/ES/04/I-04).

2. CELL LINES

In this work different human and murine cell lines have been used. As transient packaging cell line for lentiviral supernatant production we used the human epithelial 293T cell line (derived from 293 cell line ATCC: CRL-1573, Rockeville, MD, USA) maintained in DMEM-Glutamax medium containing 10% FBS and 0.5% antibiotics (50 U/mL penicillin and 50 µg/mL streptomycin; P/S, Gibco Life Technologies) at 3-5x10^5 cells/mL. This cell line was originally obtained from human embryo kidney in which the temperature sensitive T antigen from SV40 virus was inserted. T antigen serves as a proto-oncogene capable of transforming 293 cells and increasing the transfection efficacy because it allows the episomal replication of those plasmids containing the replication origin and promoter regions from SV40. For titration of the viral stocks produced by transfection, HT1080 cell line (ATCC: CCL-121) was used. This line derives from human epithelial fibrosarcoma and was maintained in DMEM-Glutamax medium with 10% FBS and 0.5% P/S at a 10^5 cells/mL.

To check the functionality of the first vector (pRRL-CMV-hRPK), the ectopic expression of hRPK transgene was detected in mouse embryo fibroblast cell line NIH/3T3 (ATCC: CRL-1658) infected with the mentioned vector and maintained in DMEM-Glutamax medium containing 10% FBS and 0.5% P/S at 3x10^5 cells/mL. To study the specificity of the vectors developed with erythroid-tissue specific promoters, three human cell lines were used, each of them with different degree of commitment towards different hematopoietic lineages. Human myeloid HL60 cell line (ATCC: CCL-240) from promyelocytic leukemia was used as a non-erythroid control. As human erythroid cell lines we used K562 (ECACC: 89121407) (Lozzio and Lozzio, 1975) and HEL (ATCC: TIB-180) (Martin and Papayannopoulou, 1982), which were originally isolated from patients with chronic myelogenous leukemia and erythroleukemia respectively. HL60 and K562 cells were maintained in IMDM-Glutamax containing 20% FBS and 0.5% P/S at 3-5x10^5 cell/mL, while HEL cells were cultured in RPMI-Glutamax medium (Gibco Life Technologies) with 10% FBS and 0.5% P/S at densities below 10^5 cells/mL. All cell lines were grown in suspension and no differentiation protocols were applied.

3. Hematological parameters study

All animals underwent a complete analysis of hematological parameters in peripheral blood, including erythrocyte and leukocyte counts, hemoglobin levels (HGB), hematocrit percentage (HTC), mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH). Fresh complete blood was collected from tail veins in 20% Ethylendiamintetracetic acid (EDTA 0.5 M, pH 8) and analyzed in an automated blood cell analyser (Abacus Junior vet.).

The percentage of reticulocytes was determined in the different groups of animals by flow cytometry using the Acridine Orange dye (stock solution 10 mg/mL in water, Invitrogen Life Technologies). This product is a cell-permeant nucleic acid binding dye that emits green fluorescence when bound to dsDNA (λ 525 nm) and red fluorescence when bound to ssDNA or RNA (λ 575 nm). This characteristic makes acridine orange useful for identifying the reticulocyte population because they have higher amounts of RNA than mature erythrocytes. Five µL of anticoagulated blood were diluted 1:400 in phosphate-buffered saline and stained with Acridine Orange at a final concentration of 0.5 µg/mL. Samples were incubated for 30 min at room temperature and directly analyzed in an EPICS XL flow cytometer (Beckman Coulter Electronics, Hialeah, FL, USA). A minimum of 5.10^5 red cells were analyzed, identifying the reticulocytes as RNA positive cells within the total erythrocyte population.

4. Study of red blood cell survival in peripheral blood

RBC survival was evaluated using biotinylation of the entire RBC cohort and in both control and transplanted mice. *In vivo* staining of RBC was carried out by three consecutive intravenous injections of 250 µL of Biotin 3-sulfo-N-hydroxysuccinimide ester sodium salt (50 mg/kg, Sigma Aldrich B5161) diluted to 4 mg/mL in saline solution. 5 µL of tail vein blood (containing approximately 3.10^7 RBC) were harvested in 50 µL PBE (0.1% BSA, EDTA 2mM in

PBS) and labelled with anti-mouse Ter119-phycoerythrin primary antibody (Ter119-PE; BD Bioscience) at 2 µg/mL to identify the RBC population, and streptavidin-fluorescein isothiocyanate (SAV-FITC; BD Bioscience) at 50 µg/mL, which specifically binds to biotin molecules. After 30 min incubation at 4ºC, cells were washed in PBA (0.1% BSA p/v, 0.02% $Na\text{-}N_3$ in PBS), centrifuged at 1400 rpm for 10 min, and resuspended in PBA with 2 µg/mL propidium iodide (IP) to exclude dead cells. The percentage of biotinylated cells within the total RBC population was monitored every 2-4 days for 40 days after the injection in an EPICS XL flow cytometer (Beckman Coulter).

5. Plasma Epo levels

Plasma Epo levels were measured in the different groups of transplanted animals. We quantified Epo concentration using the Quantikine Mouse Epo immunoassay kit (MEP00B R&D Systems) that employs the quantitative sandwich enzyme immunoassay technique. Standards, control and samples diluted 1:2 were pipetted into the wells pre-coated with an immobilized monoclonal antibody specific for mouse Epo and incubated for 2 hours at room temperature. After washing away any unbound substances, a horseradish peroxidase (HRP)-linked monoclonal antibody against mouse Epo was added to the wells and incubated for 2 hours at room temperature. Then, tetramethylbenzidine (TMB) substrate solution was added to the wells and a blue colour developed in proportion to the amount of Epo present in each sample. Colour development was stopped with the stop solution from the kit turning the colour in the wells to yellow. The absorbance of the colour was measured at 450 nm in a Genious Pro Tecan reader. Epo concentration in the plasma of the different mice was calculated by interpolating the samples absorbance in the standard curve provided by the manufacturer.

6. STRUCTURAL AND HISTOLOGICAL STUDIES

Presence of splenomegaly was evaluated in animals from either characterization studies or gene therapy experiments. Spleens were collected, photographed and weighed on precision scales. Histological studies were performed on spleens and livers harvested from the different animal cohorts, fixed in 70% ethanol (Scharlau Chemie, Barcelona, Spain), dehydrated in a crescent series of ethanol solutions and embedded in paraffin (Merck). Sections of 5 µm thickness were prepared and mounted on glass slides, followed by deparaffinization in xylene (Scharlau Chemie, Barcelona, Spain) and rehydration in decreasing ethanol solutions of 100%, 95%, 70%, 50% (v/v) and finally in water. The sections were then stained with hematoxylin (Gill-2 Haematoxylin, Thermo, Pittsburg, USA) and eosin (Eosin Alcoholic, Thermo), dehydrated in ethanol 95%, 100 %, fixed in HistoClear (National Diagnostics, Hessle Hull, UK) and mounted in DPX (Thermo Scientific). Liver slides were separately stained with Prussian Blue or Perls' staining for iron deposits detection (Sigma Aldrich, HT20). Prussian Blue reaction involved the treatment of sections with acid solutions of ferrocyanide, which combines with the ferric iron (Fe^{3+}) present in the tissues resulting in the formation of a bright blue pigment. Liver sections were incubated with a solution of potassium ferrocyanide solution 4% p/v and hydrochloric acid solution 1.2 mmol/L mixed in 1:1 proportion for 10 min at room temperature. After rinsing with distilled water, sections were stained with pararosaniline solution 1% p/v in methanol for 3-5 min at room temperature, rinsed with distilled water and dehydrated and mounted as described above. All sections were examined using an Olympus BX40 light microscope and photographed with an Olympus DP21 camera using a final magnification of 20x or 40x.

7. STUDY OF HEMATOPOIETIC PROGENITORS

7.1 Study of hematopoietic lineages in peripheral blood

Peripheral blood from the tail vein of transplanted animals was harvested, and 50 µL were lysed using cold ammonium chloride buffer (0.155 mM NH_4Cl, 0.01 mM $KHCO_3$, 0.1 µM

EDTA) for 10 min at room temperature. Samples were then centrifuged at 1400 rpm at 4 ºC for 10 min, washed in PBA and centrifuged again. Cells were then labelled for the myeloid lineage using the biotynilated antibodies Gr1 (BD Bioscience) and Mac1 (BD Bioscience) both used at 5 µg/mL, incubated for 30 min at 4 ºC and washed with PBA before centrifuging to harvest the cells. Samples were then stained for lymphoid markers using CD3-PE, B220-PE and B220-PECy5 antibodies (BD Bioscience) at 10 µg/mL, as well as SAV-TRC secondary antibody (Invitrogen) for 30 min at 4 ºC (Table 4). Therefore, the following lineages could be identified:

- T-lymphocytes: CD3-PE$^+$ cells
- B- lymphocytes: double B220-PE$^+$ and B220-PECy5$^+$ cells
- Granulocytes and monocytes: Gr1-Biot$^+$ and Mac1-Biot$^+$ cells

After the second incubation, cells were washed in PBA, centrifuged at 1400 rpm for 10 min and resuspended in PBA containing DAPI (4',6-diamidino-2-phenylindole) 2 µg/mL (Boehringer, Ingelheim, Germany) to remove dead cells. Samples were analysed in a BD LSR Fortessa Cytometer (Bekton Dickinson, Palo Alto, CA, USA).

7.2 Study of erythroid differentiation

Erythroid differentiation was studied using the combination of the erythroid-specific marker Ter119 and the differentiation marker CD71, since it is well established that the expression of transferrin receptor (CD71) on erythroid cells is required for iron overtake and HGB synthesis (Chang et al., 2004). The analysis of the expression intensities of both markers allows the identification of four murine erythroblasts populations at different maturation stages as shown in figure 10.

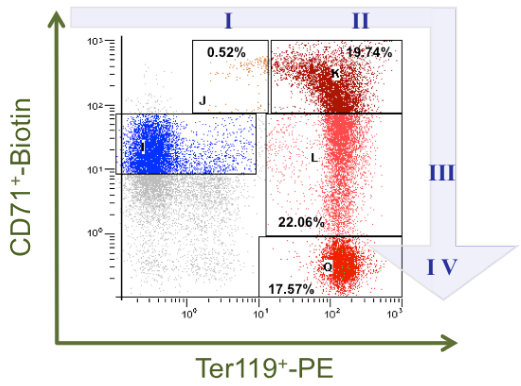

Figure 10: Study of erythroid differentiation. Flow cytometry strategy used to identify the different erythroid subpopulations according to the expression intensity of Ter119 and CD71 surface markers.

I: early proerythroblasts

II: basophilic erythroblasts

III: late basophilic and polychromatophilic erythroblasts

IV: orthochromatophilic erythroblasts and mature erythroid cells

Flow cytometry analysis was performed on BM, spleen and liver samples from the different groups of animals. Bone marrow cells were harvested by flushing two femurs and two tibiae per mouse with Iscove modified Dulbecco medium (IMDM Glutamax-1 Gibco Life Technologies). Spleen cells were obtained using the GentleMACS™ dissociator (MACS Miltenyi Biotec) to homogenize the organ in 5mL of PBE. Cell clumps were removed by passing the suspension through a 100 µm cell strainer (BD Falcon). Livers from different mice were carefully broken up into small pieces in 5 mL of Hanks' Balanced Salt Solution (HBSS Gibco Invitrogen) containing 0.25% of DNase I (2 mg/mL, Roche) and collagenase A (1 mg/mL, Roche) to avoid cell aggregation, incubated for 10 min at 37 ºC and homogenized in the GentleMACS™ dissociator (Miltenyi Biotec). The suspension was then passed through a 100 µm cell strainer, washed with HBSS and centrifuged at 1200 rpm for 5 min at room temperature. Approximately 10^6 cells from BM, spleen or liver were stained with anti-mouse PE-conjugated Ter119 (BD Bioscience) antibody at 4 µg/mL and biotinylated anti-CD71 (BD Bioscience) at 10 µg/mL at 4 ºC for 30 min (Table 4). Then, cells were washed with PBA, centrifuged at 1400 rpm for 10 min and incubated with streptavidin-tricolor (Invitrogen) at 4 ºC for 30 min. Finally, cells were resuspended in PBA with 2 µg/mL IP and analyzed in an EPICS XL flow cytometer (Beckman Coulter) to quantify the percentage of each erythroid subpopulation described above in the text.

7.3 Study and characterization of leukaemia development

Bone marrow, spleen and peripheral blood from those transplanted mice that showed aberrant expression of the hematopoietic lineages markers were harvested, and a panel of antibodies was used in order to characterize the leukaemia by flow cytometry. Lysed samples from the three organs were labelled with a cocktail for lymphoid markers containing anti-mouse FITC-Nk1.1 (8 µg/mL, BD Bioscience), PE-CD4, PECy5-CD3 (3.2 µg/mL, BD Bioscience) and Biotin-CD8 antibodies, and stained with SAV-PECy7 secondary antibody (Caltag) (Table 4). We also labelled the samples with a cocktail for myeloid markers: FITC-Gr1 (8 µg/mL BD Bioscience), PE-Mac1, PECy5-B220 (3.2 µg/mL, BD Bioscience) and PECy7-cKit (3.2 µg/mL, BioLegend), and non-lysed samples were stained with PE-Ter119 antibody (3.2 µg/mL, BD Bioscience) for erythroid lineage plus FITC-CD71 (8 µg/m, BD Bioscience) and PECy7-CD41 (3.2 µg/mL, eBioscience) antibodies. Samples were then analysed in a BD LSR Fortessa Cytometer (Bekton Dickinson).

8. Hematopoietic progenitor cultures

Bone marrow from different cohorts of mice was harvested by flushing the femurs and tibiae under sterile conditions with IMDM Glutamax-1 (Gibco Life Technologies) containing 20% of foetal bovine serum (FBS, Gibco Invitrogen), washed twice with non-supplemented IMDM and resuspended in a suitable volume of IMDM. Spleen cells were obtained using the GentleMACS™ dissociator as explained before. Cells were then washed twice with IMDM and resuspended in a suitable volume of IMDM. Either $2x10^4$ to $5x10^4$ cells from BM or $3x10^5$ to $5x10^5$ cells from spleen were resuspended in 1 mL of semisolid methylcellulose-based medium (MethoCult GF M3434; StemCell Technologies, Vancouver, BC, Canada) and seeded in a 25-mm^2 Petri dish. Each sample was measured in triplicate. This medium contains 1% methylcellulose in IMDM, 15% FBS, 1% bovine serum albumin (BSA), 200 µg/mL human transferrin (iron-saturated), 10 µg/mL human recombinant insulin, 1 mM 2-mercaptoethanol, 2 mM L-glutamine and recombinant growth factors such as murine stem cell factor (rmSCF, 50 ng/mL), murine interleukin 3 (rmIL-3, 10 ng/mL), human interleukin 6 (rhIL-6, 10 ng/mL) and human erythropoietin (rhEpo, 3U/mL) to support growth of the differently committed progenitors. After 7 days of incubation at 37ºC in a 5% CO_2 and 95% relative humidity environment, colony-forming units (CFUs; cluster of 30 or more cells) were scored using an inverse phase contrast Nikon ELWD 0.3 microscope. Results were expressed as the average of the total number of CFUs per 10^5 seeded cells in each dish.

9. Mobilization capacity of hematopoietic progenitors

The mobilization of hematopoietic progenitors was studied by injection of a single dose of pegylated recombinant human granulocyte colony stimulating factor (PEG rhG-CSF 10 mg/mL, Neulasta, Amgen, Breda, Holland) into 10-15 week-old PKD and C57BL/5 mice. PEG rhG-SCF was injected subcutaneously in a single dose of 50 µg per mouse (2.5 µg/g) diluted in 200 µL of PBS with 1% BSA. In parallel, control mice from both strains were injected with PBS 1% BSA as vehicle controls. Four days after injection, mice were sacrificed

and 1 mL of peripheral blood was collected and lysed with buffered ammonium chloride (0.155 mM NH_4Cl, 0.01 mM $KHCO_3$, 10^{-4} mM EDTA) in sterile conditions for 15 min at room temperature and washed with PBA. CFU assays were performed according to standard techniques. $2x10^5$ or $9x10^4$ mononuclear cells were seeded in triplicate on 1 mL semisolid methylcellulose-based medium (MethoCult GF M3534; StemCell Technologies) containing rmSCF (50 ng/mL), rmIL-3 (10 ng/mL), rhIL-6 (10 ng/mL) and rmGM-CSF (0.01 ng/mL), as well as 1% methylcellulose in IMDM, 15% FBS, 1% BSA, 200 µg/mL human transferrin (iron-saturated), 10 µg/mL human recombinant insulin, 1 mM 2-mercaptoethanol and 2 mM L-glutamine. Cultures were maintained at 37ºC, 5% CO_2 and 95% relative humidity for 7 days to subsequently count the number of CFUs in an inverse phase contrast Nikon ELWD 0.3 microscope.

10. Study of the hematopoietic stem cell content (LSK and SLAM populations)

Bone marrow and spleen from 15-25 week-old PKD, C57BL/6 and AB6F1 mice were obtained and processed as described in section 8. Cellularity was calculated by counting the number of nucleated cells in a light microscope using Turk solution (2% acetic acid, 0.01% methylene blue in water) and expressed as the number of cells per spleen or per bone in the case of BM (total nº of cells divided by three, considering 1 femur = 2 tibiae). BM and spleen suspensions were lysed using cold ammonium chloride buffer for 10 min at room temperature and washed with PBA. LSK stem population was identified by labelling 10^6 BM cells and $2x10^6$ splenocytes with a lineage depletion cocktail containing anti-mouse biotin-conjugated antibodies (B220, Ter119, CD3, Gr1 and Mac1, DB Bioscience) at 5 µg/mL, and subsequent incubation with 2.5 mg/mL SAV-TRC (BD Bioscience) secondary antibody (Table 4). Cells were then stained with anti-mouse FITC-conjugated Sca1 (8 µg/mL, DB Bioscience) and PE-conjugated cKit (3 µg/mL, BD Bioscience) antibodies, and analysed in a BD LSR Fortessa Cytometer (Bekton Dickinson, Palo Alto, CA, USA). Separately, Lin^- cells were labelled with antibodies against the SLAM family of surface markers: PE-conjugated CD150 and APC-conjugated CD48 (2 µg/mL, eBioscience) as an alternative for HSC identification.

MONOCLONAL ANTIBODIES

Antigen	Clone	Conjugated	Isotype	Concentration (mg/mL)	Manufacturer
B220 (CD45R)	RA3-6B2	Biotin	Mouse $IgG_{2a,κ}$	0.5	BD Bioscience
B220 (CD45R)	RA3-6B2	PE	Rat $IgG_{2a,κ}$	0.2	BD Bioscience
B220 (CD45R)	RA3-6B2	PECy5	Rat $IgG_{2a,κ}$	0.2	BD Bioscience
CD150 (SlamF1, IPO-3)	9D1	PE	Rat IgG_1	0.2	eBioscience
CD3e	145-2C11	Biotin	Hamster IgG	0.5	BD Bioscience
CD3e	145-2C11	PE	Hamster IgG	0.2	BD Bioscience
CD3e	145-2C11	PE-Cy5	Armenian hamster $IgG2_κ$	0,2	BD Bioscience
CD4	H129,19	PE	Rat $IgG_{2a,κ}$	0,2	BD Bioscience
CD41	NWReg30	PECy7	Rat $IgG_{1,κ}$	0,2	eBioscience
CD48 (SlamF2, BCM1)	HM48-1	APC	Armenian hamster	0.2	eBioscience
CD71	C2	FITC	Rat $IgG_{1,κ}$	0,5	BD Bioscience
CD71	C2	Biotin	Rat $IgG_{1,κ}$	0.5	BD Bioscience
CD8 (Ly-2)	53-6,7	Biotin	Rat IgG_{2a}	0,5	BD Bioscience
cKit (CD117)	2B8	PECy7	Rat $IgG_{2b,κ}$	0,2	BioLegend
cKit (CD117)	2B8	PE	Rat $IgG_{2b,κ}$	0.2	
Gr1 (Ly6G, Ly6C)	RB6-8C5	FITC	Rat $IgG_{2b,κ}$	0,5	BD Bioscience
Gr1 (Ly6G, Ly6C)	RB6-8C5	PE	Rat $IgG_{2b,κ}$	0.2	BD Bioscience
Gr1 (Ly6G, Ly6C)	RB6-8C5	Biotin	Rat $IgG_{2b,κ}$	0.5	BD Bioscience
Mac1 (CD11b)	M1/70	PE	Rat IgG_{2b}	0,2	BD Bioscience
Mac1 (CD11b)	M1/70	Biotin	Rat $IgG_{2b,κ}$	0.5	BD Bioscience
Nk1.1	PK136	FITC	Mouse (C3HxBalb/c) $IgG2_{a,κ}$	0,5	BD Bioscience
Sca1 (Ly-6A/E)	E13-161.7	FITC	Rat $IgG_{2a,κ}$	0.5	BD Bioscience
Ter119 (Ly76)	Ter119	Biotin	Rat $IgG_{2b,κ}$	0.5	BD Bioscience
Ter119 (Ly76)	Ter119	PE	Rat $IgG_{2b,κ}$	0.2	BD Bioscience

SECONDARY ANTIBODIES and DYES

Product	Concentration (mg/mL)	Provider
Acridine Orange	10 µg/mL	Invitrogen Detection Technologies, Eugene, Oregon, USA, A3568
DAPI	2 µg/mL	Boehringer, Ingelheim, Germany
IP	2 µg/mL	BD Bioscience 556463
SAV-FITC	0.5 mg/mL	BD Bioscience 554060
SAV-PECy7		Caltag SA1012
SAV-TRC		Invitrogen SA1006

Table 4: Flow cytometry antibodies used in the different experiments. Table indicates the antibody name, the clone where was produced, the conjugated fluorocrome, the isotype, and the concentration used, as well as the manufacturer company.

11. SELF-INACTIVATING LENTIVIRAL VECTORS

11.1 Construction of lentiviral plasmids carrying ubiquitous promoters

In this work two different ubiquitous promoters were used to drive the expression of the RPK transgene: the cytomegalovirus strong promoter (CMV), and the relatively weak eukaryotic promoter of the human PGK gene. For the generation of vectors we used lentiviral backbones originally derived from pRRL.sin18.ppt.CMV-EGFP.Wpre and pCCL.sin.ppt.hPGK-EGFP-Wpre*constructs, kindly provided by Dr. Naldini (HSR-TIGET, San Raffaele Telethon Institute for Gene Therapy and Vita Salute San Raffaele University Medical School, Milano, Italy).

11.1.1 Development of pRRL-CMV-hRPK lentiviral vector

The first set of vectors harbouring the CMV promoter were obtained using the pRRL.sin18.ppt.CMV.EGFP.Wpre HIV-derived lentiviral backbone (7425 bp) which carries the Rous Sarcoma Virus (RSV) promoter regions in the 5' LTR, and the CMV as the internal promoter (≈599 bp) driving the expression of the enhanced green fluorescent protein (EGFP) (Figure 11). This construction was used as control vector, and the pRRL.CMV-hRPK therapeutic vector (7798 bp; Figure 11) was obtained by inserting the 1731 bp fragment corresponding to the open reading frame of the human *PKLR* gene (PK EC 2.7.1.40, Gene Bank ID 5313) into the same lentiviral backbone via XbaI + EcoRI digestion (New England BioLabs).

In this first design, pRRL.CMV-EGFP-Wpre control lentiviral constructions carried the post-transcriptional regulatory element Wpre, which has been seen to increase the transgene expression in a lentiviral context (Zufferey et al., 1999). However, it contains a truncated version of the WHV X gene involved in animal liver cancer (Kingsman et al., 2005), so we decided not to use it in the vector carrying the therapeutic hRPK transgene to avoid safety concerns. A fragment of 3895 bp covering the expression cassette (cPPT-CMV-hRPK)

was sequenced to check that no harmful mutation occured during the cloning process, using primers designed with VectorNTI software (Invitrogen) and detailed in table 5.

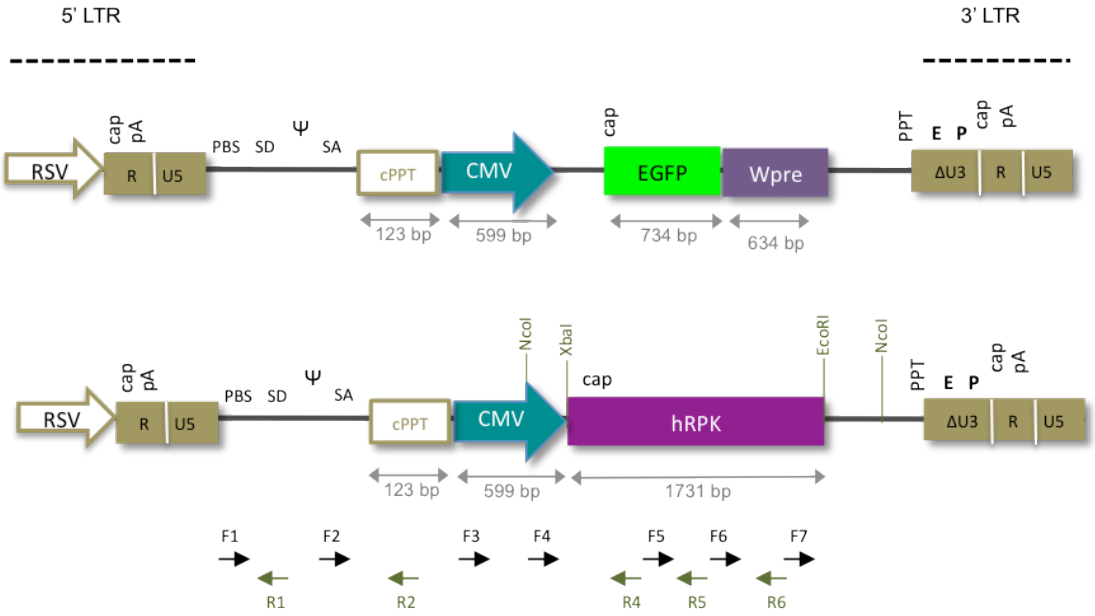

Figure 11: Schematic representation of the pRRL-CMV-hRPK and pRRL-CMV-EGFP-Wpre lentiviral constructions.

Primer	5'	Sequence	3'	% GC	Tm (ºC)
F1		CCATTAGGAGTAGCACCCAC		55	59.4
F2		CGCAAAACCAGCAAGAAAAG		45	56
F3		CGATAAGCTTGGGAGTTCCGC		57.1	61.8
F4		CCACGCTGTTTTGACCTCCA		55	59.4
F5		CCAGAGTCGGAAGTGGAGCT		60	61.4
F6		GAGGTTTGATGAAATCCTGG		45	59.4
F7		CGACCTCGGGCAGCAGTCATT		61.9	63.7
R1		CCAGGTCGTGTGATTCCAAA		50	55.3
R2		CCGTCGAGATCCGTTCACTA		55	59.4
R4		GGACTCAGCATGGTACTCGT		55	59.4
R5		CGTCGCTGGCTTTCCGCACAAA		59.1	64
R6		GGATCACGGCTTAGTGGCGCT		61.9	63.7

Table 5: Primers used for sequencing of the pRRL-CMV-hRPK lentiviral construction.

11.1.2 Development of pCCL-PGK-coRPK-Wpre* lentiviral vector

To clone the therapeutic vector bearing the human PGK ubiquitous promoter and the *PKLR* cDNA transgene we used a different strategy, which was based on the *in silico*

optimization of the transgene design to achieve stable and increased protein expression. The optimized version of the *PKLR* gene (coRPK) was kindly designed by Dr. Axel Schambach (Hannover Medical School, Germany) using GeneArt® software. Optimized RPK sequence was designed with a high GC content, and algorithms to avoid cryptic splice sites were also applied. AgeI and SacII restriction sites were inserted flanking the *PKLR* optimized sequence in 5' and 3' respectively and the Kozak consensus sequence was also included in the transgene, due to its important role in the initiation of the translation process (Kozak, 1996). The coRPK optimized sequence showed 80.4% homology with the human *PKLR* cDNA and 76.5% with the mouse *Pklr* cDNA, with no changes in the amino acid sequence of the protein (Appendix 1). Once we had the coRPK sequence, a synthetic DNA fragment of 1743 bp was produced by GeneScript Corporation (Piscataway, NJ, USA) and inserted in a pUC57 plasmid.

In a second step, polymerase chain reaction (PCR) was performed using 50 ng of pUC57-coRPK plasmid as template and primers A and B at 0.4 µM, which included two additional restriction sites and the first or last 20 bp of the coRPK transgene (Figure 12). Amplification was carried out using 5 units of the highly thermostable Pfx50 polymerase (Invitrogen), 0.4 mM of each dNTP and 1.2 mM Mg^{2+} in the PCR buffer. Samples were incubated at 94 ºC for 2 min and underwent 34 cycles of 94 ºC/30 sec; 60 ºC/ 30 sec; 68 ºC/2 min and a final incubation of 68 ºC/2 min.

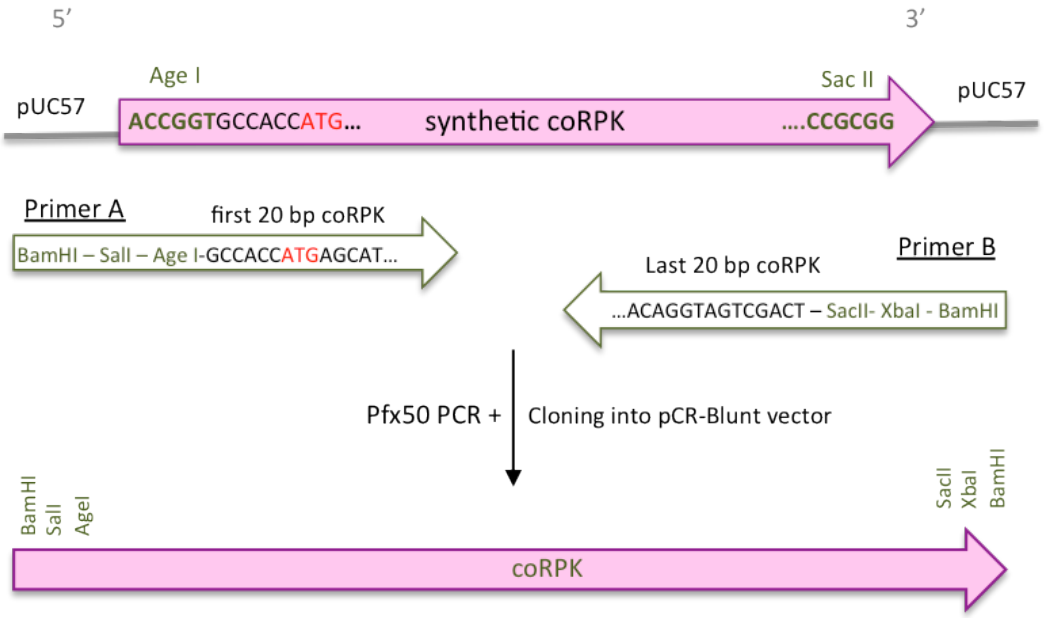

Figure 12: Design and synthesis of the *PKLR* optimized version. The figure shows a schematic representation of the synthetized coRPK transgene and the primers used to obtain a suitable product for cloning into pCR-Blunt vector. Primers used for sequencing the resulting vector are also shown.

A 1761 bp blunt ended dsDNA product was obtained and inserted into a pCR-Blunt cloning vector (Zero Blunt cloning kit, Invitrogen Life Technologies). Positive *E.coli* One Shot Top10 transformants (Invitrogen Life Technologies) were confirmed by restriction analysis and PCR. Amplification was performed using primers C and D (detailed in table 6), 0.2 mM of each dNTP, 0.3 µM of each primer, 1.75 mM $MgSO_4$ and 5 units of *Taq* DNA polymerase (Invitrogen) using the following conditions: 94 ºC/5min and 30 cycles of 94 ºC/30 sec, 58 ºC/30 sec and 72 ºC/45 sec; and a final incubation of 5 min/72 ºC. To ensure the proper synthesis and correct orientation of the coRPK insert, we next sequenced the pCR-Blunt-coRPK plasmid using universal M13 and primers C to G detailed in table 6.

Primer	5' Sequence 3'	% GC	Tm (ºC)
A	GGATCCGTCGACACCGGTGCCACCATGAGCAT	62.5	79
B	GGATCCTCTAGACCGCGGTCAGCTGATGGACA	59.4	78
C	CGGATACCTGAGAAGGGCTA	55	60
D	GCTGCTGCTGGAAAAAGG	55.6	56
E	GCTGAACTTCTCCCACGGCA	60	63
F	CTACATCGACGACGGCCTGA	60	63
G	AATGATGATCGGACGGTGCA	50	58
H	ATCTTCGCCGCTACCCTTGT	55	60
I	ATTCTGCAAGCCTCCGGAGC	60	63
J	GGATGACCAGATAGTATGAA	60	63
K	CTCGAGGATATCATCGATAATTCGATTTTGGGTTCAAA	36.8	72
K	GTCGACACCGGTTCACTAGTGATTGAGCCCCA	56.3	76
EGFP	GTAAACGGCCACAAGTTCAGC	52.4	61
EGFP	TGGTGCAGATGAACTTCAGGG	52.4	61
S1	GCGCGCAATTAACCCTCACT	55	60
S2	TACGGTAAACTGCCCACTTG	50	58
S3	GAGTCCTGCGTCGAGAGAGC	65	65
S4	CGAACAGGGACCTGAAAGCG	60	63
S5	AACCCTCTATTGTGTGCATC	45	56
S6	TGCTGAGGGCTATTGAGGCG	60	63
S7	TGCTGGTTTTGCGATTCTTC	45	56
S8	GGATGGAGTGGGACAGAGAA	55	60
S9	TTAGTGAACGGATCTCGACG	50	58
S10	GAGCGCACCATCTTCTTCAA	50	58
S11	CCAGGATTTATACAAGGAGG	45	56
S12	GGCTTTCGTTTTCTCCTCCT	50	58
S13	CCAATGACTTACAAGGCAGC	50	58
S14	GTTGTAAAACGACGGCCAGT	50	58
S15	GGCAGCGGCTACACCAACAT	60	63
S16	CTGGGCACTGATAATTCCGT	50	58
S17	CCTCTTCGCTATTACGCCAG	55	60
S18	TGTAACGCGGAACTCCATAT	45	56

Table 6: Primers used for cloning and sequencing the pCR-Blunt-coRPK and pCR-Blunt-PKLR plasmids and pCCL.PKG.coRPK, pCCL-PKLR-coRPK and pCCL-PKLR-EGFP lentiviral constructions. Forward primers are written in black font, while reverse ones are in green font.

To develop a lentiviral vector carrying the coRPK transgene and the human PGK ubiquitous promoter we used a pCCL lentiviral backbone carrying the CMV promoter in the 5' LTR region, because previous results pointed to a higher viral titer than the pRRL vectors. We also used the mutated version of the Wpre sequence, which has a mutation in the open reading frame of the WHV *X* gene (Wpre*) (Zanta-Boussif et al., 2009), avoiding safety concerns. For this cloning, the 1761 bp coRPK fragment was inserted into pCCL.sin.ppt.hPGK-Wpre* lentiviral construction via BamHI digestion. pCCL-PGK-coRPK-Wpre* (8942 bp) final construction was checked by both PCR with C and D primers (as described above in the text) and BamHI restriction analysis. Additional double digestions were applied to verify the correct orientation of the insert. A fragment of 6 Kb covering most of the vector sequence was subsequently sequenced using primers gathered in figure 13 and table 6, to rule out any possible mutation affecting the expression of the transgene.

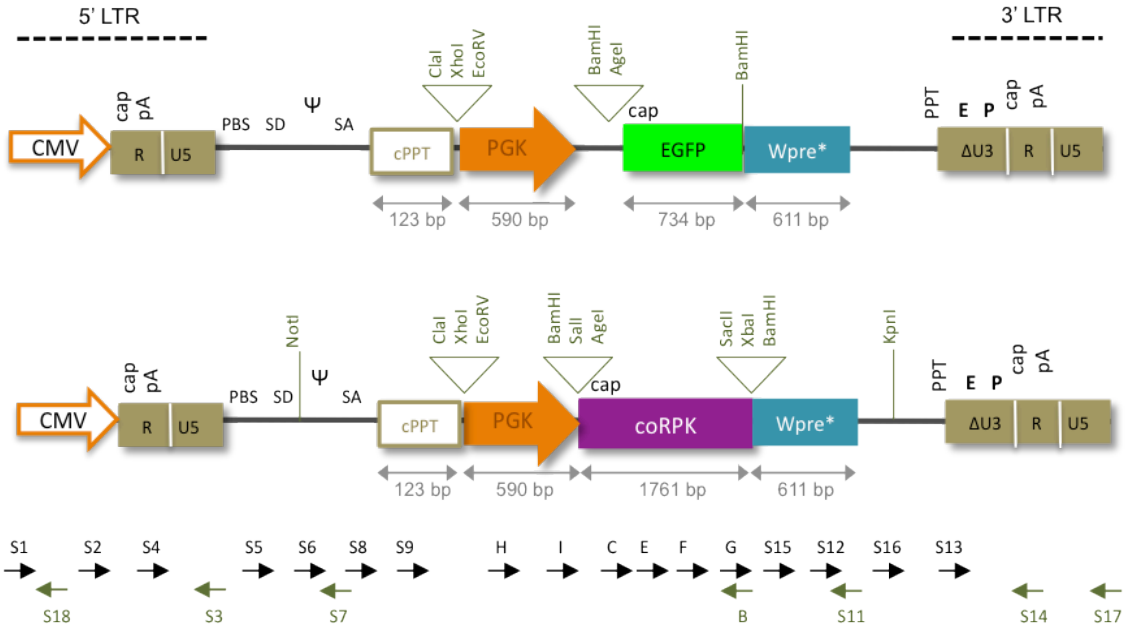

Figure 13: Schematic representation of the pCCL-PGK-EGFP-Wpre* and pCCL-PGK-coRPK-Wpre* lentiviral constructions. The upper part of the figure shows the pCCL-PGK-EGFP-Wpre* control vector, which also served as the starting point for obtaining the pCCL-PGK-coRPK-Wpre* vector (bottom figure), in which the analyzed restrictions sites are represented. Black (forward) and green (reverse) arrows represent the primers used for vector sequencing.

11.2 Construction of lentiviral plasmids carrying erythroid-specific promoters

In addition to PGK and CMV ubiquitous promoter, regulatory regions of the human *PKLR* gene and human β globin promoter (βp) were used to develop erythroid-specific lentiviral vectors in order to restrict the transgene expression to the erythroid lineage. However, tissue-specific promoters are usually inefficient to reach therapeutic levels, so we also included different combinations of DNase I HS from the human βLCR.

11.2.1 Development of pCCL-PKLR-coRPK-Wpre* lentiviral vector

The development of a physiologically regulated vector with the coRPK transgene expression regulated by its own promoter (PKLR) was performed using a PCR based strategy. The PKLR promoter sequence used to develop these vectors corresponds to a 601 bp fragment containing the 507 bp upstream of the human *PKLR* gene transcription start site and the first 94 bp of its open reading frame. The primer pair was designed containing the restrictions sites needed to substitute PGK by PKLR promoter in the pCCL-PGK-coRPK-Wpre* construction previously developed (Figure 14).

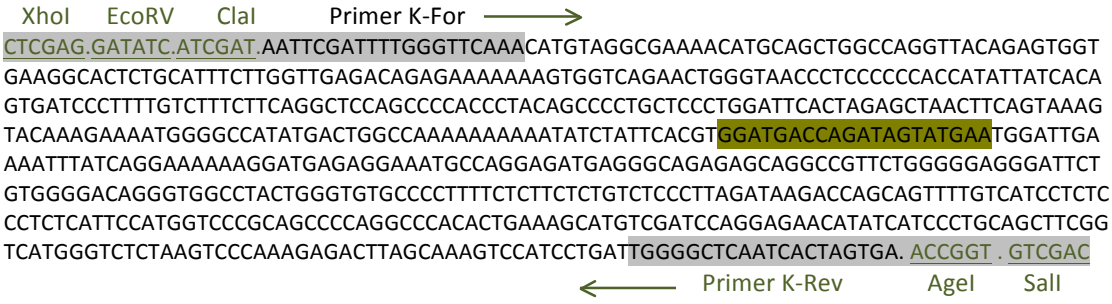

Figure 14: Sequence of the PKLR promoter. The figure shows the sequence of the PKLR promoter used to develop the erythroid-specific lentiviral vector carrying the coRPK transgene. The restriction sites added by PCR are highlighted, as well as primers K-For, K-Rev (in grey) and J (in green).

A 650 bp fragment was synthetized the same way as described in *section 11.1.B* (primers A and B) using 50 ng of pGEMT-PKLR as template and Pfx50 polymerase. The resulting blunt product (PKLR promoter) was next cloned in a pCR-Blunt vector and positive transformed *E.coli* One Shot Top10 clones were confirmed by PCR using M13Rev and J primers also detailed in figure14. We also checked the introduced restriction sites by double

digestion and sequencing using M13-For, M13-Rev and J primers. XhoI-SalI PKLR fragment was cloned into pCCL-coRPK-Wpre* lentiviral construction in which PGK had been previously removed. Positive clones for pCCL-PKLR-coRPK-Wpre* (9041 bp) (Figure 15) were checked first by PCR using J and D primers and by double digestion. In addition, pCCL-PKLR-coRPK-Wpre* final lentiviral construction was sequenced using the primers shown in figure 15 and table 6, covering a total region of 6 kb. Finally, to have an appropriate control vector for further experiments, we replaced coRPK for the EGFP transgene via AgeI and BamHI double digestion obtaining the pCCL-PKLR-EGFP-Wpre* construction (8022 bp) (Figure 15). After transformation, positive clones were checked by both PCR, using primers EGFP-For and EGFP-Rev (same conditions as for C and D primers, see section 11.1.B), and double digestion.

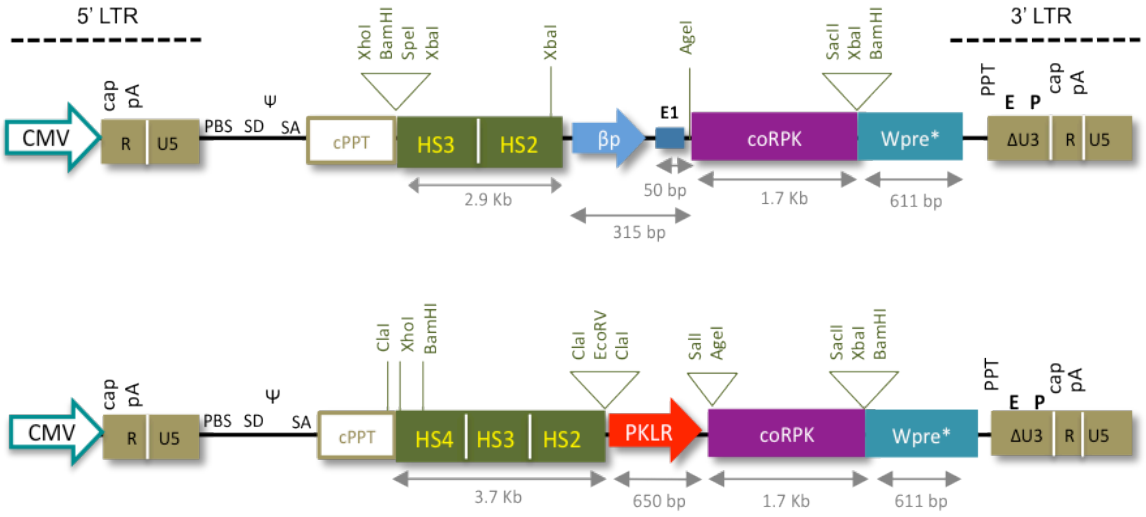

Figure 15: Schematic representation of the pCCL-PKLR-EGFP-Wpre* and pCCL-PKLR-coRPK-Wpre* lentiviral constructions. In the figure both restriction sites used for cloning and primers used for sequencing are represented.

11.2.2 Development of erythroid-specific lentiviral vectors carrying the human β-globin LCR sequences

To avoid the lack of efficient expression associated with the use of tissue-specific promoters, we developed two new lentiviral vectors harbouring either the PKLR or the β-globin (βp) erythroid-specific promoters, incorporating several HS sequences from the human β-globin LCR, generously provided by Dr. Antoniou (School of Medicine, Guy's

Hospital London, King's College London). For the first cloning, AgeI-XhoI 2.9 kb fragment containing the HS3/HS2/βp sequences was inserted into pCCL-PKLR-coRPK-Wpre* lentiviral vector previously digested at AgeI-XhoI sites for removing the 650 bp PKLR promoter. pCCL-HS3-HS2-βp-coRPK-Wpre* final lentiviral construction (11.3 Kb) was then checked by restriction analysis after *E.coli* One Shot Top10 competent cell transformation. The second developed vector, however, carried the PKLR promoter and the additional HS4 sequence together with HS3 and HS2 to enhance the promoter activity. It was obtained by ligation of the EcoRV-HS4-HS3-HS2-XhoI 3.7 Kb fragment with the linearized pCCL-PKLR-coRPK-Wpre* vector at the same restriction sites. The final pCCL-HS4-HS3-HS2-PKLR-coRPK-Wpre* (12.7 Kb) lentiviral construction (Figure 16) was checked by restriction analysis and *E.coli* One Shot Top10 competent cells were transformed in order to amplify the plasmid DNA.

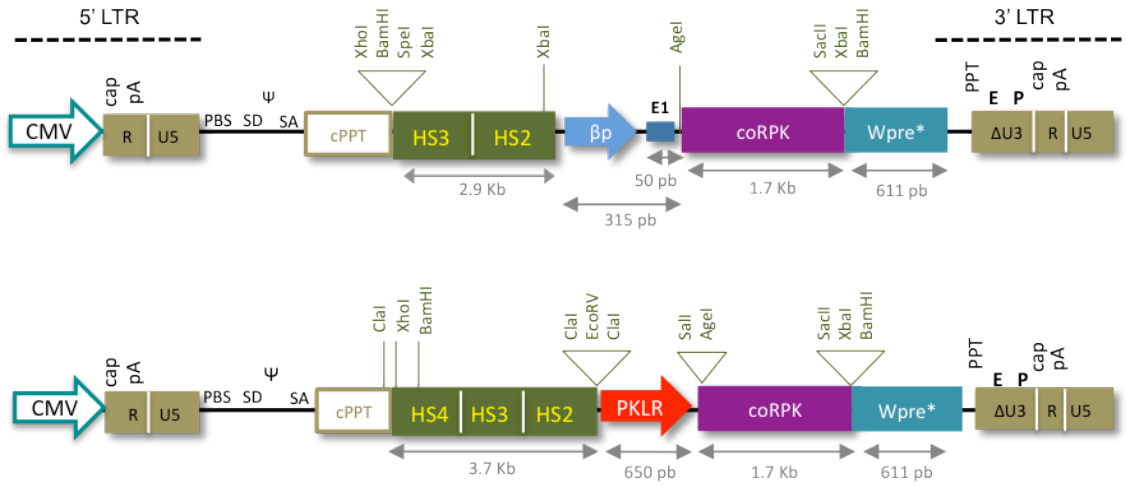

Figure 16: Schematic representation of enhanced erythroid-tissue specific lentiviral constructions. Restriction sites used for cloning are indicated in both HS3-HS2-β Globin promoter-coRPK (upper part of the figure) and the physiologically regulated vector HS4-HS3-HS2-PKLR-coRPK (lower part of the figure).

12. LENTIVIRAL VECTORS PRODUCTION

All self-inactivating HIV-1-derived vectors generated as we described above were produced by transient cotransfection of 293T cells, obtaining VSV-G-pseudotyped viruses that were next titrated to determine the quality of the produced lentiviral supernatants. Transfections were performed on 293T cells at 70-80 % confluence in 150 mm diameter

plates and lentiviral supernatant productions were carried out following the $CaCl_2$ DNA precipitation method previously described elsewhere (Dull et al., 1998); (Vigna et al., 2005). Briefly, one hour before transfection culture medium was replaced by IMDM-Glutamax containing 10% of FBS and antibiotics (50 U/mL penicillin and 50 µg/mL streptomycin; P/S, Gibco Life Technologies). Mixtures of three plasmids containing the vector genome and the packaging constructs that encode only the proteins essential for LV assembly and functions were prepared freshly. 293T cells were co-transfected with 40 µg of the gene transfer plasmid (pRRL or pCCL lentiviral vectors) carrying the transgenes and containing the viral genome, 16.25 µg of the pCMVdR 8.74 packaging plasmid (Plasmid Factory) carrying the *gag-pol-rev* viral genes, and 7 µg of the pMD2-VSV-G envelope plasmid (Plasmid Factory) carrying the heterologous envelope VSV-G per plate. These plasmid mixtures were prepared in a final volume of 1125 µL of 0.1x TE/dH_2O (2:1) and 125 µL of 2.5M $CaCl_2$ were carefully added. After a 5 min incubation at room temperature, 1250 µL of 2x HBS buffer (HEPES 100 mM (GibcoBLR), NaCl 281 mM, Na_2HPO_4 1.5 mM, pH 7.13) was added drop by drop, allowing the formation of Ca^{2+}-DNA^{-} precipitates that would be subsequently phagocyted by the cells. 14 hours after, medium-containing precipitates was replaced by fresh medium and transfection efficacy was measured at 24h post-transfection by flow cytometry in cells transfected with EGFP harbouring vectors. Lentiviral supernatants were collected 24 and 48 h post-transfection, filtered through 0.22 µm pore-size filter (Millex GP, Milipore) to eliminate cell debris and concentrated approximately 300-fold by ultracentrifugation (25 000 rpm, 2h, 4ºC). Viral pellets were then resuspended in PBS, aliquoted and stored at -80ºC.

13. TITRATION OF LENTIVIRAL SUPERNATANTS

Vector titers were determined by EGFP expression analysis (in vectors carrying the fluorescent transgene) or quantitative PCR (qPCR) on genomic DNA (gDNA) of the infected cells. HT1080 cells were seeded in 24-well tissue culture plates at 50 000 cells/well. The next day, cells were infected with serial dilutions of each vector stock (from 10^{-2} to 10^{-6}) in a total volume of 200 uL of DMEM-Glutamax, 10% FBS, 0.5% P/S. Twenty four hours after infection, viral supernatants were removed and replaced by fresh medium. Infected cells were maintained in culture for 7 to 11 days before harvesting to avoid episomal DNA detection.

For the quantification of viral titers in coRPK lentivirus supernatants, integrated proviral DNA copies in H1080 gDNA were determined 11 days after infection. Proviral genome was amplified using a 7500 Fast Real-Time PCR System (Applied Biosystems) at the Ψ packaging sequence by multiplex qPCR following the *Taq*MAN method, and normalized to the human albumin gene (hAlb) intron 12 as previously described (Charrier et al., 2005). For the detection of the Ψ lentiviral sequence, the primers and probe used were as follows: Ψ-For 5'-CAGGACTCGGCTTGCTGAAG (Tm 63 ºC), Ψ-Rev 5'-TCCCCCGCTTAATACTGACG (Tm 59 ºC) and Ψ-probe 5'-6FAM-CGCACGGCAAGAGGCGAGG-BHQ1 (Tm 75.7 ºC). For the detection of the hAlb housekeeping sequence we used the following primers and probe: hAlb-For 5'-GCTGTCATCTCTTGTGGGCTG (Tm 64 ºC) and hAlb-Rev 5'-ACTCATGGGAGCTGCTGGTTC (Tm 63 ºC) primers and hAlb-Probe 5'-TxRed-CCTGTCATGCCCACACAAATCTCTCC-BHQ2 (Tm 73.9 ºC). Amplification reactions (20 µL) contained 50-100 ng of sample gDNA and 10 µL of *Taq*MAN buffer (TaqMAN Universal PCR Master Mix, No AmpErase UNG, Applied Biosystems), 0.1 µM primers (forward and reverse) and 0.1 µM TaqMAN probe and consisted of a 50 ºC/2 min holding, 95 ºC/10 min and 40 cycles of 95 ºC/15 sec and 60 ºC/1 min.

Quantification of the number of Ψ and hAlb molecules present in the sample was carried out by preparation of standard curves with ten-fold dilutions of pRRL-SIN.cPPT-PGK-EGFP-hAlb plasmid (7539 bp) and known amount of DNA (starting at 10^6 molecules). This lentiviral plasmid was provided by Dr Charrier (Genethon, Evry, France) and contained the relevant sequences (Ψ and hAlb). The number of viral genomes or particles per transduced cell (VCN/cell) in the samples was determined after interpolation into the standard curve (Figure 17) and normalization of Ψ with hAlb, assuming two alleles hAlb per cell and two Ψ copies per integration (VCN/cell = $[VCN_{Psi} \cdot 2] / VCN_{hAlb}$). Titers expressed in vp/mL were calculated as:

Titer (vp/mL) = normalized VCN/cell · nº infected cells · supernatant dilution factor

All VCN/cell values obtained were corrected using in parallel the HT1080 cell clone HT4 developed by (Charrier et al., 2011) which carries a single integration of a rHIV lentiviral provirus verified by Vector Integration Tag analysis (VITA), qPCR and southern blot. Measures were performed at least in duplicate and non-transduced HT1080 cultured in

parallel were used as negative controls. Data were analysed using the 7500 software v2.0.1 (Applied Biosystems).

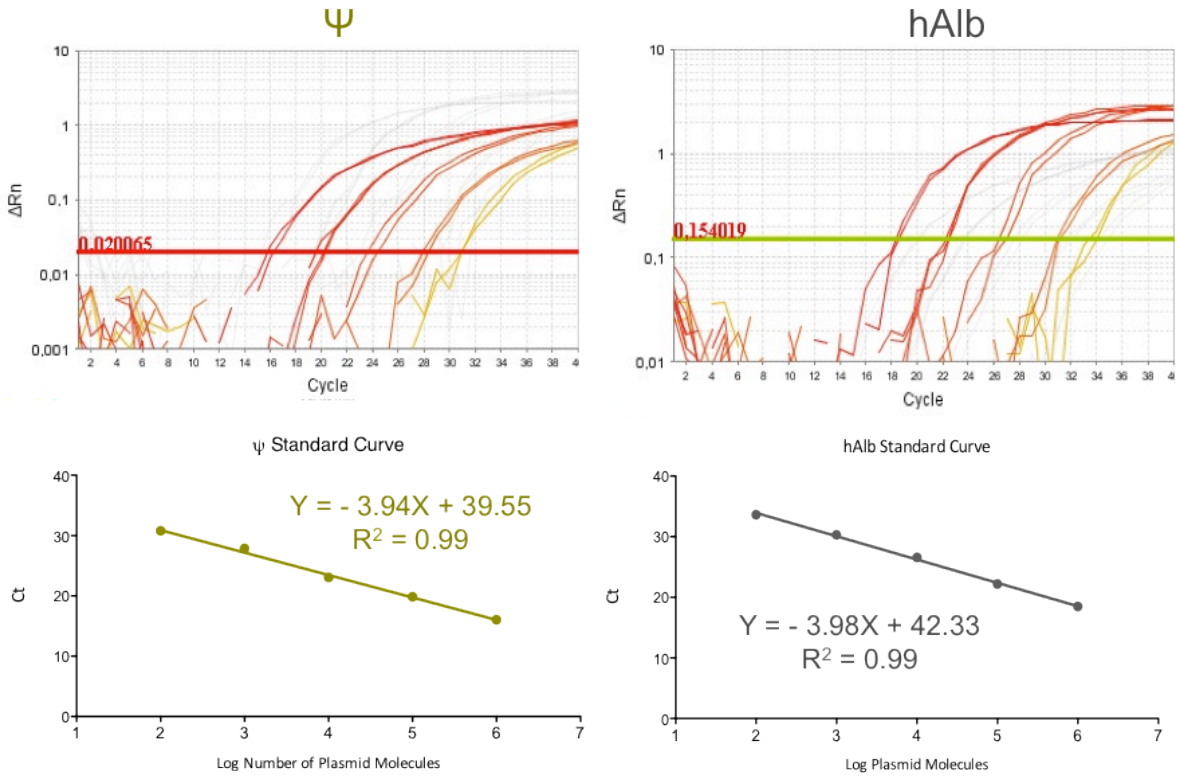

Figure 17: pRRL-SIN.cPPT-PGK-EGFP-hAlb plasmid standard curves. The upper part of the figure shows the amplification plots for each detected gene and its threshold cycle. In the lower part the standard curves for each gene are represented, showing the equation obtained by linear regression and the R^2 coefficient.

Additionally, EGFP lentiviruses were titrated by harvesting transduced cells 7 days after infection and analysing EGFP expression in an EPICS XL flow cytometer. Viral titers were calculated according to the following formula:

$$\text{Titer (vp/mL)} = \frac{\text{Nº infected cells} \cdot \%\ \text{EGFP}^{+}\ \text{cells}}{100} \cdot \text{CF} \cdot \text{DF}$$

where CF is the correction factor to express the titer in vp/mL (CF = 1000 µL / infection volume in µL) and DF is the vector supernatant dilution factor in each sample. With this method the viral titer was calculated as infective particles per mL (ip/mL) because it

measures the functional viral particles able to express the EGFP transgene. However, viral titers obtained by qPCR were measured as total viral particles per mL (vp/mL) considering both functional and non-functional viral particles. We also observed that titers obtained by qPCR method were one logarithm higher than the ones obtained by FACS analysis for the same EGFP vector, confirming the results from Sastry et al (Sastry et al., 2002) probably because of the higher sensitivity of the technique. In order to homogenize the multiplicity of infection in the transduction protocols we considerer all the viral titers obtained by qPCR subtracting one logarithm.

14. Isolation and transduction of mouse bone marrow hematopoietic progenitors (Lin- cells)

Bone marrow from 8-14 week-old male PKD mice was harvested in sterile conditions by flushing the two leg bones with 0.5 mm x 16 mm needles and non-supplemented IMDM-Glutamax. Lineage negative cells (Lin^-) (the BM fraction enriched in hematopoietic progenitors) were purified using the Lin^- Cell Depletion kit (Miltenyi Biotec, Gladbach, Germany) based on the magnetic separation of mature hematopoietic cells by staining with biotinylated antibodies against CD5, B220 (CD45R), Mac1, Gr1 and Ter119 for 30 min at 4ºC. Both positive and negative BM fractions were collected and lineage negative cell purity was evaluated by flow cytometry using B220, Gr1, Mac1, Ter119 and CD3 PE-conjugated monoclonal antibodies, obtaining purities of 70-90%.

Lin^- cells were pre-stimulated by culture in IMDM-Glutamax medium supplemented with 20% FBS, 0.5% P/S and mSCF and hIL-11 growth factors (100 ng/mL) at 37 ºC and 5% CO_2 for one day. Pre-stimulated cells ($4x10^5$ – $4x10^6$ cells/mL) were then transduced with the EGFP or hRPK/coRPK expressing lentiviral vectors, with titers ranging from 10^7 to 10^8 vp/mL. Depending on the experiment, we performed one to three consecutive infections at MOI 1-20. Each infection was carried out for 24h at 37ºC in the presence of the cytokines mentioned above on 6-well tissue culture plates previously incubated overnight with the CH-296 fibronectin fragment (2 µg/cm^2; Retronectin, TakaraShuzo, Otsu, Japan) at 4ºC.

15. Mouse hematopoietic progenitors transplantation

Cells transduced with the different lentiviral vectors were harvested 24 hours after infection, carefully resuspended in PBS and centrifuged at 1200 rpm for 10 min. Proliferation and viability of the cells were examined by the Tripan blue test. $5x10^4$ to 10^6 transduced hematopoietic cells per mouse were intravenously injected into the tail vein of 8-14 week-old female PKD mice, which had been lethally irradiated with 9 to 9.5 Gy in two split doses of 4.5 Gy or 4.75 Gy respectively, 24h apart using a Philips MG324 X-ray instrument (Philips, Hamburg, Germany). Transplanted mice were monitored at different time-points for 3-6 months. Secondary transplants were performed at the end-point of each experiment using the female receptors from the primary transplants as donor mice. Total BM was harvested from the primary receptors, carefully washed with PBS and centrifuged at 1400 rpm for 10 min. The number of nucleated cells was determined by counting the BM cell suspensions with Turk reagent for either freezing the cells in viable conditions or freshly transplanting into lethally irradiated 8-14 week-old female PKD secondary receptors. Transplants were performed as explained before by injecting $4x10^6$ total BM cells per mouse, from either frozen vials or freshly harvested cells. Secondary receptor mice were analysed at different time-points as previously mentioned.

16. Transduction of cell lines with self-inactivating lentiviral vectors

For 3T3 cell transduction, 10^5 cells were transduced with pRRL-CMV-EGFP or pRRL-CMV-hRPK lentiviral vectors by four consecutive 12h infections at 10 vp/cell. Cells were maintained in culture for 8 days, pelleted by centrifugation (1300 rpm/10min) and stored at -80ºC to further analyse coRPK protein expression by western blot.

For the erythroid-specificity studies HL60, K562 and HEL cell lines were transduced with different lentiviral vectors harbouring either ubiquitous (PGK) or erythroid-specific promoters (PKLR, HS3-HS2-βp or HS4-HS3-HS2-PKLR) and expressing the EGFP or the coRPK

transgene. Approximately $8x10^4$ cells from each line were infected in suspension with the different lentiviral vectors with titers ranging from $5x10^6$ to 10^8 vp/mL. Transduction was carried out at $4x10^5$ cells /mL in a volume of 200 µL and a MOI of 4 vp/cell. 24h later, fresh medium was added to the cells to dilute the viral supernatant tripling the final volume. Transduced cells were maintained for 2-3 weeks assaying cell growth and viability by the Tripan blue dye exclusion test. Ten days after the infection, EGFP expression was analysed in an EPICS XL flow cytometer in cells transduced with control vectors, while cells transduced with coRPK harbouring vectors were harvested at 12-20 days after infection, washed with fresh medium and pelleted by centrifugation (1300 rpm/10min). Pellets were then store at -80ºC to further analyse coRPK protein expression.

17. Molecular studies

17.1 Vector copy number calculation

The number of proviral integrations per cell was analysed in transplanted mice by qPCR. Total BM and peripheral blood samples (100 µL) were collected and lysed with ammonium chloride buffer (0.155 mM NH_4Cl, 0.01 mM $KHCO_3$, 0.1 µM EDTA). Genomic DNA from nucleated cells was isolated using the DNeasy Blood & Tissue kit (Qiagen) and 20-50 ng of sample gDNA were amplified by qPCR using detection of Ψ proviral sequence normalized to the mouse *Titin* (TTN) housekeeping gene, as described for titrations (*Section 13*). Titin primers and probe used were as follows: TTN-For 5'-AAAACGAGCAGTGACGTGAGC (Tm 61ºC), TTN-Rev 5'-TTCAGTCATGCTGGTAGCGC (Tm 60 ºC) and TTN-Probe 5' TxRed-TGCACGGAAGCGTCTCGTCTCAGTC-BHQ2 (Tm 76.3 ºC). Reactions were performed in 20 µL containing 10 µL of *Taq*MAN buffer, 0.1 µM primers (forward and reverse) and 0.1 µM TaqMAN probe, and consisted on 50 ºC/2 min, 95 ºC/10 min and 40 cycles of 95 ºC/15 sec and 58 ºC/1 min.

Standard amplification curves were obtained by serial dilutions of the pRRL-hGPK-EGFP-Wpre-Titin lentiviral construction (7584 bp) kindly provided by Dr Charrier (Genethon, Evry, France), which contained the appropriate sequences from the vector and the TTN

gene. Linear regression coefficient of the standard curve was accepted above 0.98 and the threshold of detection was determined with this standard. Amplified DNA, simultaneously detecting the Ψ sequence and the TTN gene, was equimolar to the amount of pRRL-hGPK-EGFP-Wpre-Titin plasmid used in each sample from the standard curve. Cells from non-transplanted mice were used as negative controls. Samples and standards were amplified at least in duplicate and analysed using the 7500 software v2.0.1 (Applied Biosystems).

17.2 Calculation of VCN in individual hematopoietic colony-forming units (CFUs)

Bone Marrow CFUs from the transplanted animals (see section 8) were well isolated by aspiration with a pipet tip under the microscope and VCN was quantified following the method described by (Charrier et al., 2011). Cells were suspended in 200 µL of PBS and each tube containing a single colony was centrifuged at 1500 rpm for 10 min. Pellets from each CFU were resuspended in 10 µL of PBS and gDNA was extracted by incubation with 20 µL of lysis buffer (0.3 mM Tris HCl, pH 7.5; 0.6 mM $CaCl_2$, 1.5 % Glycerol; 0.675 % Tween-20 and 0.3 mg/mL Proteinase K) at 65 ºC/30 min, 90 ºC/10 min and 4ºC/10 min. After lysis, 10 µL of gDNA preparation were amplified by qPCR in a total volume of 25 µL and in triplicate, the same way as explained in section 17.1, also using negative controls (CFUs from non-transplanted animals). 11 to 33 individual CFUs per mice were analyzed, where VCNs below 0.1 copies per cell were considered negative.

17.3 Donor chimerism analysis

Quantification of donor chimerism was accomplished by multiplex qPCR using the 7500 Fast Real-Time PCR System (Applied Biosystems) as previously described (Navarro et al., 2006). Genomic DNA from peripheral blood of transplanted mice was isolated and the SRY sequence of the Y chromosome and the β-Actin genomic sequences were amplified to quantify male donor cells in female receptor mice. Percentages of Y-chimerism were calculated by reference to standard curves consisting of varying proportions from male and female blood gDNA. Primers for the male-specific sequences were as follows: SRY-For 5'-TGTTCAGCCCTACAGCCACA (Tm 53.9 ºC) and SRY-Rev 5'-CCTCTCACCACGGGACCAC (Tm 54.8 ºC) and detected with the TaqMAN probe SRY-P 5'-6FAM-ACAATTGTCTAGAGAGCATGGAGGGCCA-

BHQ1 (Tm 64.7 ºC). The murine genomic β-Actin sequences were amplified using the primers mβAct-For 5'-ACGGCCAGGTCATCACTATTG (Tm 53.9 ºC) and mβAct-Rev 5'-ACTATGGCCTCAGGAGTTTTGTCA and detected with the TaqMAN probe mβAct-P 5' TxRed-AACGAGCGGTTCCGATGCCCT-BHQ2 (Tm 63.5 ºC). Reactions were carried out in 20 µL containing 15 ng of gDNA, 10 µL of *Taq*MAN buffer, 0.1 µM primers (forward and reverse) and 0.1 µM TaqMAN probe. The thermal profile was one hold of 10 min at 95 ºC and 40 cycles of 95 ºC/30 sec and 58 ºC/1 min. For standard curves, 5 to 8 points with different percentage of donor cells were used, obtaining linear regression coefficients in the range of 0.97-0.99. Samples and standards were amplified in triplicate and analysed using the 7500 software v2.0.1 (Applied Biosystems). Donor engraftment percentage was calculated using the ΔCt method with a reference gene, a modified formula of the Livak method (Livak and Schmittgen, 2001), assuming efficiency 2 for both pair of primers:

$$\% \text{ Donor Engraftment} = 100 \cdot 2^{\Delta Ct} = 100 \cdot 2^{(Ct\ \beta Act - Ct\ SRY)}$$

17.4 Human RPK expression study by RT-qPCR

Total RNA was isolated from 100 µL lysed PB samples from PKD and control mice, using Trizol reagent (Invitrogen Life Technologies), which isolated the DNA by precipitation with chloroform and isopropyl alcohol. Glycogen (Roche Diagnostics, Mannheim, Germany) was also used as a carrier to improve the yield of the extraction. In all samples, 500 ng of total RNA were reverse transcribed into cDNA with the RetroScript RT kit (Ambion, Applied Biosystems) using 5 µM Oligo(dT) primers, 0.5 mM of each dNTP, 0.5 U of RNase inhibitor (40 U/µL, Ambion, Applied Biosystems) and 100 U of MMLV reverse transcriptase (100 U/µL, Ambion, Applied Biosystems). The reaction was performed by incubation at 72ºC/3 min, 4ºC/3 min and 55ºC/1h, to further amplify the cDNA by qPCR using primers that specifically bound with the human RPK sequence, but not with the murine sequence. The detection sequences were as follows: RPK4-For 5'-ATATCATCCCTGCAGCTTCG (Tm 63.9 ºC) and RPK4-Rev 5'-CAGCTCCTGGGTCAGTTGG (Tm 65.9 ºC) primers at 0.5 µM, using Power SYBR Green PCR Master Mix (Applied Biosystems) to detect amplified products. After 50 ºC/2 min and 95 ºC/10 min incubations, amplification was performed as 40 consecutive cycles of 95 ºC/25 sec and 60 ºC/1 min. Melting curves were generated at 72 ºC/10 sec, 60 ºC/1 min and 95

ºC/30 sec. For housekeeping control expression, primers to amplify the murine GAPDH gene sequence were used at 0.5 µM: mGAPDH-For 5'-TCCAAGGAGTAAGAAACCCTGGA (Tm 60 ºC) and mGAPDH-Rev 5'-GAAATTGTGAGGGAGATGCTCAG (Tm 59 ºC).

Relative expression of the target transgene was calculated in comparison to mGAPDH reference gene using a modified version of the Pfaffl method (Pfaffl, 2001). Due to the way primers were designed, only the ectopic expression of the RPK was detected, and no calibrator sample was available. Therefore, calibration curves were prepared with 10-fold diluted mixes containing the same amount of each cDNA sample, starting with 2 µg of total cDNA at 130 ng/µL. Ct cycles (Y) versus ng of cDNA input (X) were plotted to calculate the slope, and the corresponding efficiencies of both primers were calculated as: $E = 10^{-1/slope}$. Relative hRPK expression was determined by the following formula:

$$\text{hRPK Expression} = \frac{\text{Efficiency hRPK}^{-\text{Ct hRPK}}}{\text{Efficiency mGAPDH}^{-\text{Ct mGAPDH}}}$$

18. WESTERN BLOT

For the detection of the human RPK protein encoded by either the *PKLR* cDNA or its optimized version (coRPK), we used a polyclonal antibody (RPK Ab) developed by Dr Meza (Táchira Medicine School, University of Los Andes, San Cristóbal, Venezuela) by rabbit immunization with the human recombinant enzyme (Diez et al., 2005). This antibody specifically recognized the recombinant human RPK and showed no cross-reactivity toward the murine isoenzyme.

Lysates from 3T3, HL60, K562 and HEL cells transduced with the different vectors were prepared from $5x10^5$ PBS-washed cells in 300 µL lysis buffer (50mM Tris pH 7.5, 150 mM NaCl, 1% NP-40, 2mM EDTA) supplemented with 100 µg/mL PhosphoStop (PMSF) and Protease Inhibitors (Roche Diagnostics). Total protein concentration was determined using the Bradford colorimetric assay (Quick Star Bradford Protein Assay kit, BioRad) and 50 ng of protein from each sample were separated by Sodium Dodecyl Sulphate-polyacrylamide gel

electrophoresis (SDS-PAGE, XCell II Blot, Invitrogen Life Technologies) 125 V/1h using 4-12% Bis-Tris gel (XCell SureLock Mini-Cell, Invitrogen Life Technologies). A prestained protein ladder (PageRuler, Thermo Scientific) was run in parallel. Proteins were then transferred (150V/1.5h/4ºC humid transfer) onto previously blocked nitrocellulose membranes (Hybond-ECL Amersham Bioscience) and incubated with RPK primary Ab diluted 1:10 000 in blocking buffer (5% milk in Tris-PBS with 0.05% Tween-20). Antibody binding was detected by incubation with a goat anti-rabbit IgG (Abcam) conjugated to horseradish peroxidase (HRP) secondary antibody diluted 1:10 000 in blocking buffer. Membranes were next blotted with 1:4 000 diluted anti-mouse β-actin (Abcam, Cambridge, UK) monoclonal antibody as loading controls, which also recognized human β-actin protein. After a 2h incubation at room temperature, the membrane was blotted with a sheep anti-mouse IgG HRP-conjugated secondary antibody diluted 1:5 000 in blocking buffer. Detection was carried out using the SuperSignal West Pico Chemiluminescent Substrate (Pierce, Rockford, USA) and Kodak Medical X-ray Films. The relativization of hRPK protein expression to β-actin levels was carried out by densitometry analysis using ImageJ software (NIH, Bethesda, MD, USA). Human peripheral blood lysates were used as positive controls and 3T3 cells transduced with EGFP expressing lentiviral vector or non-transduced cells from the three human lines were used as negative or endogenous expression controls.

19. METABOLOMIC STUDIES

RBCs and WBCs were isolated from the three different control mice and PKD mice transplanted with either hRPK- or EGFP-expressing progenitors to determine if metabolic abnormalities could be corrected. RBCs were obtained from 200-400 µL of whole blood by incubating at 37ºC/2h and centrifuging at 2500 rpm for 10 min to remove the serum and WBC interface. WBCs were obtained by centrifugation of 500 µL non-coagulated PB. Pellets were stored at -80ºC until the time of metabolite extraction. Metabolite extraction was carried out by adding 5 volumes 70°C 70% ethanol to one volume of cell pellet. Cell extract was then centrifuged at 3000 rpm to pellet the precipitated protein. 30 µL of supernatant were transferred to a 96-well plate and dried under nitrogen. Samples were then resuspended in 30 µL of water to allow for metabolite profiling by mass spectrometry.

Two approaches were used to collect metabolite profiles from the isolated blood components: 1) a targeted triple-quadruple-liquid chromatography mass spectrophotometry (LCMS) approach for the analysis of glycolytic and tricarboxylic acid cycle components, and 2) an untargeted direct injection approach coupled to a high-mass-resolution Q-ToF. The targeted approach allowed for direct analysis of upstream/downstream metabolites from the pyruvate kinase enzyme, while the untargeted approach allowed for unbiased analysis of distal metabolic pathways. Statistically insignificant results were achieved for the targeted analysis of central carbon metabolites, while the untargeted approach showed clusters of ions that significantly changed.

20. Statistical analysis

Data from all experiments were represented as the average ± the standard error of the mean (SEM). To analyse if the differences among groups were significant, we applied two different non-parametric statistical tests depending on the number of different groups. When there were three or more experimental groups we used the Kruskal-Wallis test ($p < 0.05$) followed by the Dunn's multiple comparison post-test ($p < 0.05$), while the Mann-Whitney test was applied when only two groups were compared. RBC half-life studies were analyzed by two-way ANOVA ($p < 0.005$). All statistical analyses were performed on groups with $n > 2$ by using GraphPad Prism software, Version 5.0a.

VI. RESULTS

1. AcB55 mouse strain characterization

AcB55 mouse strain was first identified during the study of alleles involved in malaria susceptibility. They were obtained by crossing sets of recombinant congenic strains bred from malaria-susceptible A/J (A) and resistant C57BL/6 (B6) mice (Fortin et al., 2001). Malaria-resistance in AcB55 mice is controlled by a locus in chromosome 3 (*Char4*) mapped to a small congenic B6 fragment linkage to a loss-of-function mutation in pyruvate kinase gene (*Pklr*).

With the aim of choosing an appropriate mouse model of PKD, we studied the hematological status of AcB55 mice (PKD mice) and other features related to the hemolytic phenotype, such as the RBC survival in peripheral blood and splenomegaly. To have a better understanding of the PKD pathogenesis, we also studied the erythroid differentiation, the hematopoietic progenitors compartment and the mobilization capacity of the AcB55 bone marrow. However, given the lack of availability of the appropriate control mice and the hybrid genome composition of the PKD mice (87.5% from A/J and 12.5% from C57BL/6), we used three different mouse strains as healthy controls: C57BL/6, A/J and the first generation obtained from the cross of both strains (A/JxC57BL/6) F1 or AB6F1 mice. Using this strategy we attempted to distinguish characteristics caused by the *Pklr* mutation from the ones linked to AcB55 singular genetic background.

1.1 Phenotypic characterization

Peripheral blood from PKD, C57BL/6, A/J and AB6F1 animals was analyzed using an automated blood cell analyser to measure the hematological variables (Figure 18). PKD animals showed a reduction of around 40% in the number of RBCs compared with the three control mouse strains (table 7) indicating severe anemia, as well as an approximate 20% decrease in the hematocrit percentage and 25% increase in the mean cell volume (MCV), consistent with the hemolytic anemia condition. On the other hand, HGB levels were reduced by 30% in PKD mice, but the amount of HGB in the erythrocytes (MCH) was 20 % higher. We also observed a 30% decrease in the WBC counts of anemic PKD mice when comparing with C57BL/6 and AB6F1 groups, although no significant differences were found in the number of WBC between PKD and A/J mice, decreasing only by 13% in PKD mice.

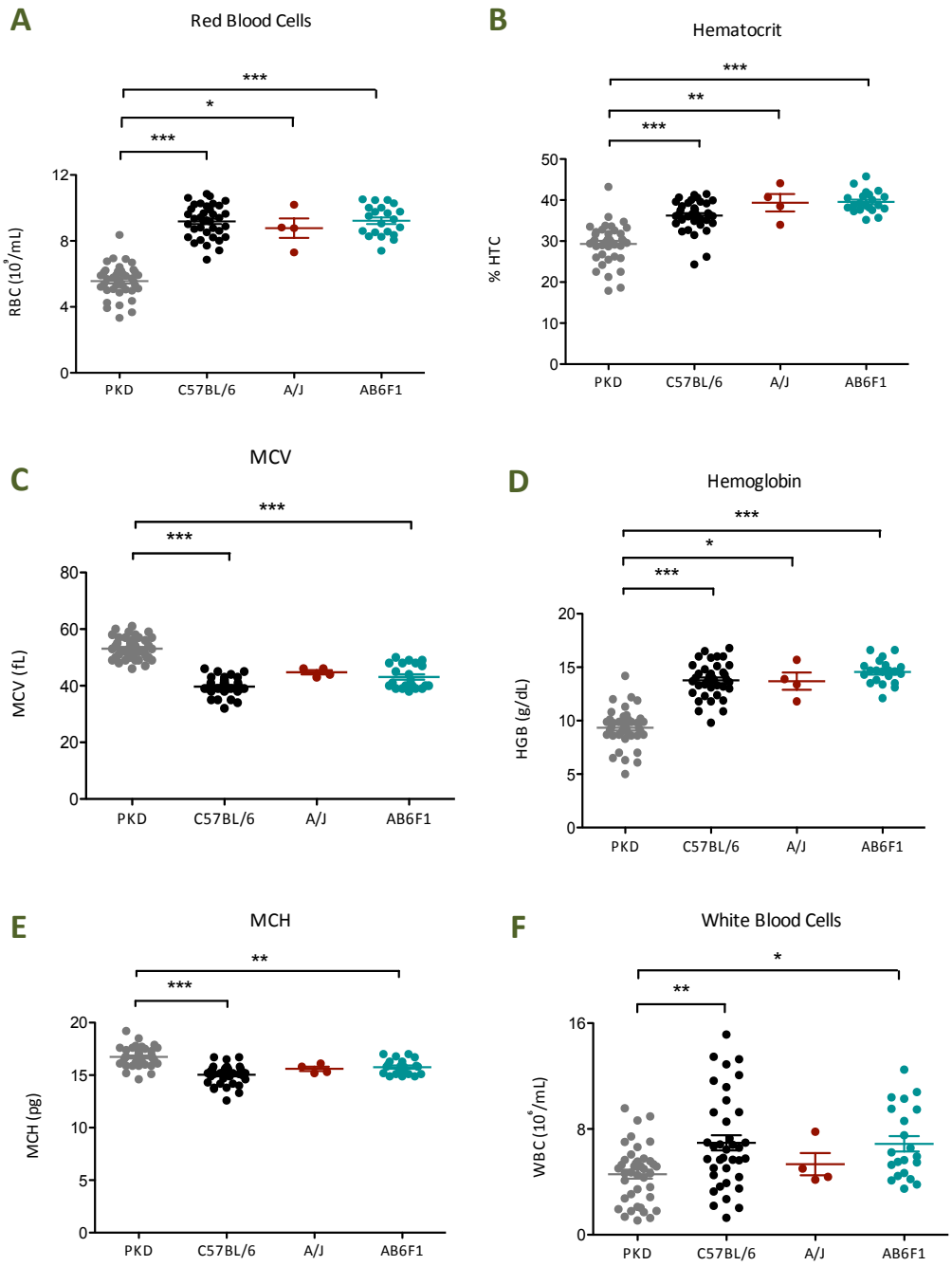

Figure 18: Hematological analysis of the different mouse strains studied. Hematological parameters recorded in peripheral blood from PKD (n=41), C57BL/6 (n= 37), A/J mice (n=4) and AB6F1 (n=21). **A)** Red blood cell counts (RBC), **B)** Hematocrit index (HTC), **C)** Mean corpuscular volume (MCV), **D)** Hemoglobin concentration (HGB), **E)** Mean corpuscular hemoglobin (MCH) and **F)** White blood cells counts (WBC). Data represent the values in each mouse, the average per group and SEM, statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$)

	decreased variables			
comparisons	**RBC**	**Hematocrit**	**HGB**	**WBC**
vs C57BL/6	60.64 ± 1.63	80.94 ± 2.02	67.85 ± 1.96	65.90 ± 4.92
vs A/J	63.45 ± 1.71	74.55 ± 1.86	68.24 ± 1.98	86.10 ± 6.42
vs AB6F1	60.35 ± 1.62	74.15 ± 1.85	64.22 ± 1.86	66.61 ± 4.97

	increased variables			
comparisons	**MCV**	**MCH**	**Reticulocytes**	**Spleen weight**
vs C57BL/6	133.48 ± 1.52	111.35 ± 0.92	645.78 ± 17.49	308.32 ± 10.71
vs A/J	118.50 ± 1.35	107.32 ± 0.89	379.04 ± 10.27	454.48 ± 15.80
vs AB6F1	123.05 ± 1.40	106.21 ± 0.88	984.09 ± 26.65	389.10 ± 13.53

Table 7: Comparisons of hematological variables between PKD mice and the three control groups. Numbers represent the % and the standard error (SEM) of PKD average (n=40) relative to each control for the different variables, considering as 100% the average of either C57BL/6 group (n=36) in the first row, A/J group (n=4) in the second row or AB6F1 (n=20) in the third row.

Flow cytometry analysis of the erythroid RNA$^+$ fraction within the total erythroid peripheral blood population allowed us to study the degree of reticulocytosis in PKD animals (Figure 19.A) because these cells have higher amounts of RNA compared to mature erythrocytes. RPK deficient mice showed a significant increase in the percentage of circulating reticulocytes, at 550% and 880% over C57BL/6 and AB6F1 levels respectively (Figure 19.B), and thereby reproducing one of the pathognomonic symptoms of PKD disease.

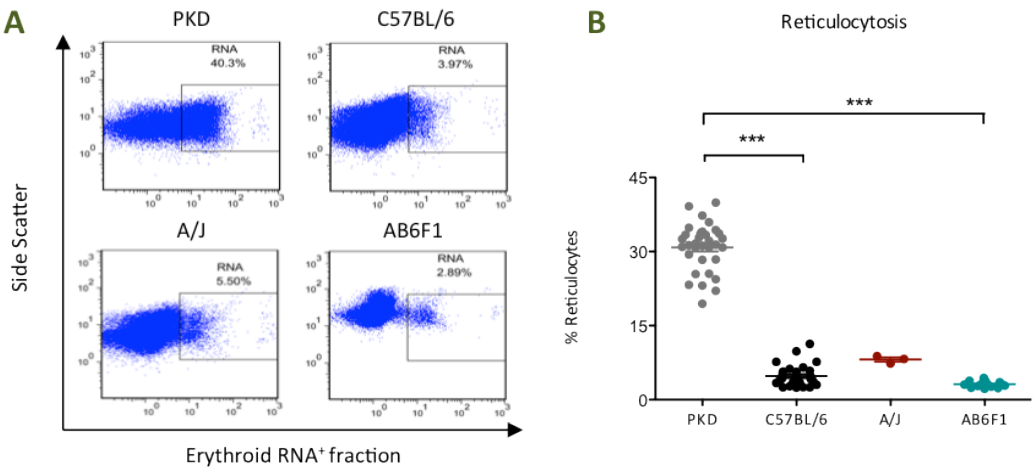

Figure 19: Study of reticulocyte levels in the different mouse strains studied. Peripheral blood from C57BL/6 (n=32), PKD (n=33), AB6F1 (n=17) and A/J mice (n=3) was analyzed. **A)** Representative dot plot of the flow cytometry analysis used to identify the reticulocyte population. **B)** Reticulocyte percentages recorded in the different groups of animals showing individual values, average and SEM in each group. Data were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

In addition, spleen size (Figure 20.A) and weight (Figure 20.B) were significantly increased in RPK deficient mice compared with controls, with an increase of 300% over C57BL/6 or AB6F1 and 350% over A/J control mice.

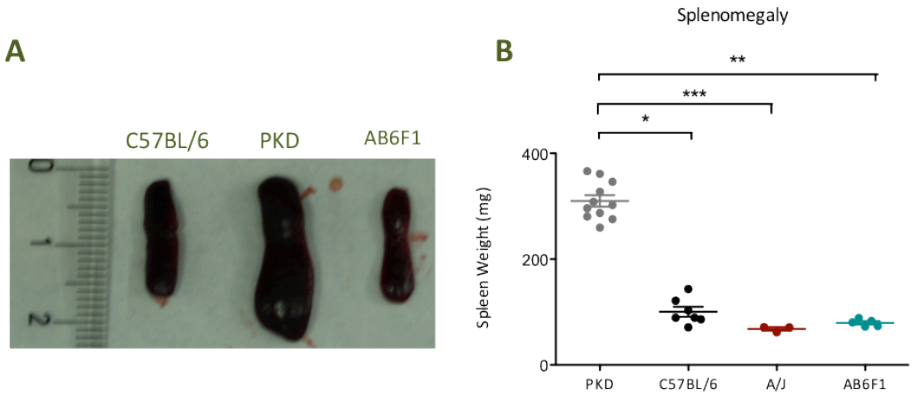

Figure 20: Study of spleen size and weight. Spleens from PKD (n=11), C57BL/6 (n=7), A/J (n=3) and AB6F1 mice (n=5) were analyzed. **A)** Pictures of representative spleens. **B)** Spleen weight from the four different groups of mice. Data represent the individual values, the average and the SEM in each group. Data were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

1.2 Study of RBC survival in peripheral blood

Because PKD animals suffer from anemia, another parameter that can be affected is the lifespan of circulating erythrocytes. We studied the survival times of RBCs in the circulation by measuring the turnover of biotin-labelled RBCs (Figure 21.A) at different time-points for 40 days after the biotin staining (*see section 4 from Materials ad Methods*). The data obtained from two independent experiments were analyzed by nonlinear regression fitting an exponential decay model, where the Y (% biotin$^+$ RBC) approaches zero at high values of X (days post-injection), which is reasonable when measuring the clearance of the labelled-RBC. Based on this model, the biotin$^+$ RBC kinetics followed the equation:

$$Y = Y_0 \cdot e^{-kx} + Plateau$$

where Y_0 is the value at time zero and *plateau* is the Y value at infinite time. In this case *plateau* tends to zero because it measures the labelled-RBC survival. K is the rate constant that determines how rapidly the curve descends.

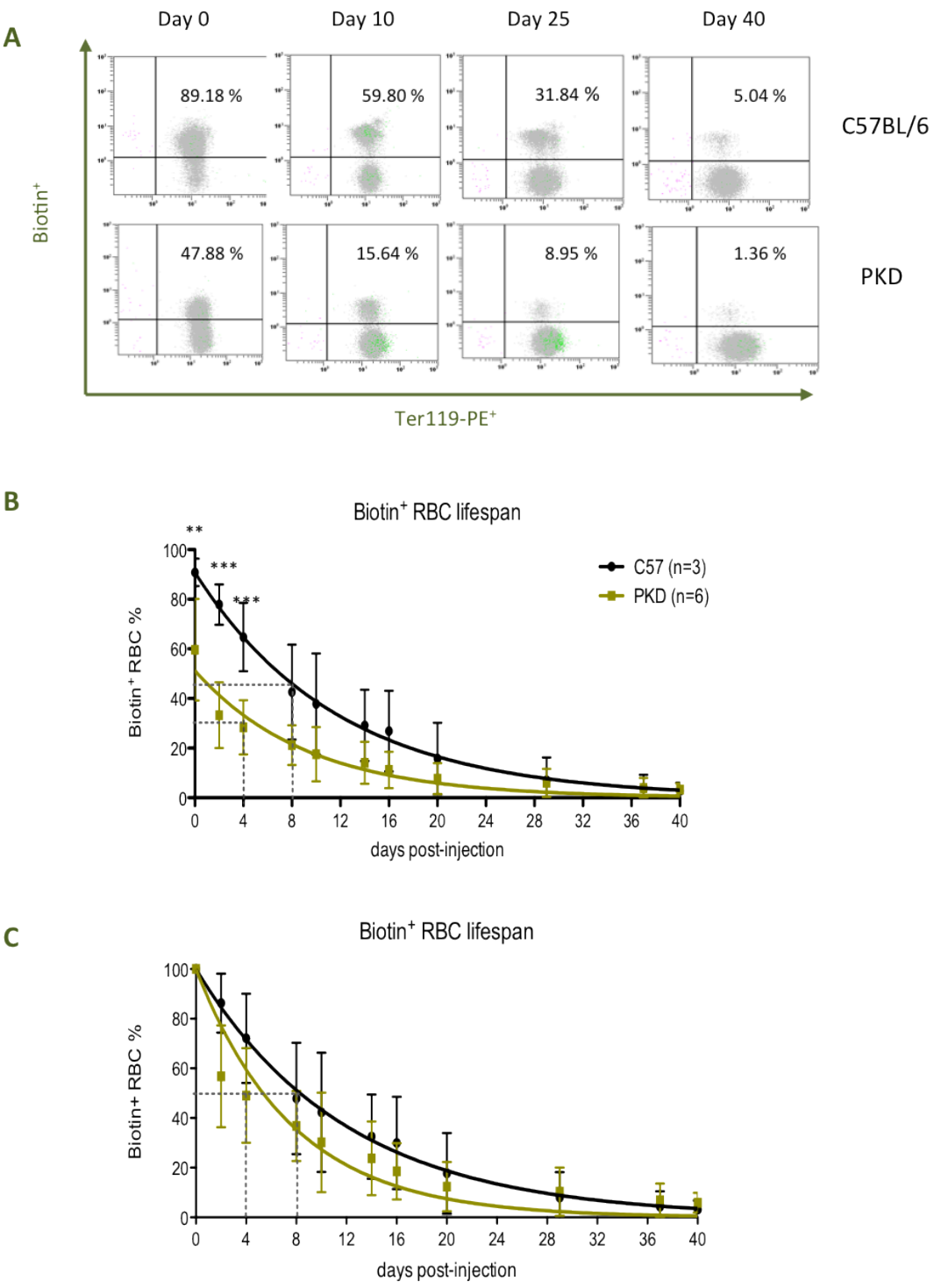

Figure 21: Study of red blood cell survival. C57BL/6 (n=3) and PKD (n=6) mice were injected with biotin 3-sulfo-N-hydroxysuccinimide ester sodium salt and peripheral blood was periodically analyzed. **A)** Representative dot plot of the flow cytometry strategy to analyse the biotinylated RBC throughout 40 days post-injection. **B)** Direct data of Biotin+ RBC percentage within the total RBC population. **C)** Relative biotin+ RBC survival within the total RBC population. The percentage obtained at day 0 was taken to be 100% for each individual mouse, and values measured throughout the time were normalized to its individual % at day 0. Data represent the average and SEM in each group from two independent experiments, and were analyzed by non linear regression fitting an exponential decay model and two-way ANOVA statistic test ($p < 0.005$).

In this assay, RBC staining was detectable up to 40 days post-injection in both types of mice, but the % of stained-RBC was significantly lower in PKD animals than in C57BL/6 mice during the first 29 days post-injection. In figure 21.B direct data are represented, showing a 40% reduction in the initial percentage of biotin$^+$ erythrocytes (Y_0) in PKD animals compared with C57BL/6 ones, as well as a higher value of K constant in the PKD group, indicating a more rapid turnover of biotin$^+$ RBC in the circulation of RPK deficient mice.

$$Y_{C57BL/6} = 90.86 \cdot e^{-0.085 \cdot x} \qquad K_{C57BL/6} = -0.085$$

$$Y_{PKD} = 51.30 \cdot e^{-0.108 \cdot x} \qquad K_{PKD} = -0.108$$

RBC half-life (Ln2/K) was also different for each data set, being 8.15 days for C57BL/6 control mice and 6.41 days for the anemic ones. Figure 21.C shows the same data relative to 100% at day 0 in each individual mouse considering Y_0 = 100%, obtaining the same results:

$$Y_{C57BL/6} = 100 \cdot e^{-0.083x} \qquad \text{half-life}_{C57BL/6} = 8.29 \text{ days}$$

$$Y_{PKD} = 100 \cdot e^{-0.129x} \qquad \text{half-life}_{PKD} = 5.34 \text{ days}$$

These results demonstrated a faster biotin clearance in peripheral blood from PKD mice compared to the control animals, showing a 1.5-fold shortened half-life of the PKD erythrocytes, as well as a reduced initial staining of RPK deficient RBCs. In general, these data agree with previously reported data (Marinkovic et al., 2007) (Min-Oo et al., 2007).

1.3 Study of Erythroid Differentiation

Differentiation patterns of the erythroid lineage were also examined, since an expansion of the erythroid compartment in the spleen of RPK deficient animals had been described (Min-Oo et al., 2004). To study the late erythropoiesis differentiation process in detail, expression of Ter119 and CD71 antigens in BM, spleen and liver from the different groups of animals were analyzed by flow cytometry. With this double staining four erythroid subpopulations can be identified, as previously described (Meza et al., 2007) (*See Materials and Methods section 7.2* Figure 10):

As shown in figure 22, a predominance of immature erythroid precursors (proerythroblasts and basophilic erythroblasts) was observed in BM and spleen from anemic

mice when comparing with C57BL/6, AB6F1 and A/J mice, being statistically significant in the earliest subpopulation (I) comparing with C57BL/6 mice. On the contrary, late erythroid subpopulation (IV) was significantly lower in RPK deficient mice than in C57BL/6 or AB6F1 groups, in both BM and spleen (Figures 22). However, this mature subpopulation of erythroid cells was similar in PKD and A/J mice, in both organs.

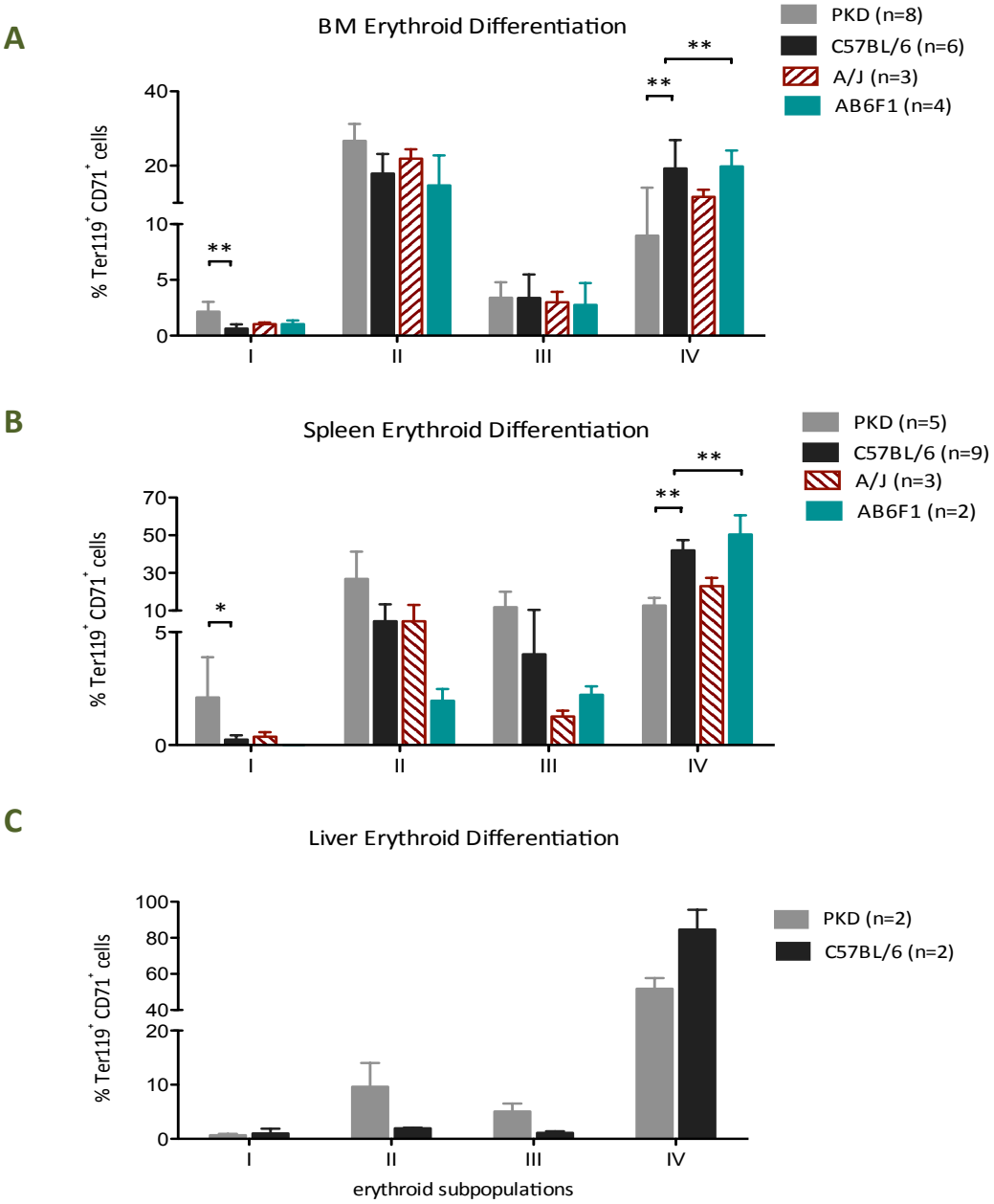

Figure 22: Study of erythroid differentiation by flow cytometry. Analysis of the different erythroid differentiation steps in **A)** bone marrow and **B)** spleen from PKD, C57BL/6, A/J and AB6F1 mice. **C)** Erythroid differentiation in livers from C57BL/6 and PKD animals. Average and SEM in each group are represented. Data were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$) when sample size was > 2.

Erythroid differentiation was also studied in livers from C57BL/6 and PKD animals (Figure 22.C), showing an increase of early subpopulations in anemic mice, mainly II and III, but a remarkable fall of late erythroid cells (IV). These data pointed out a compensatory proliferation of proerythroblast compartment in PKD animals, but this increment is not enough to offset the RBC loss caused by hemolytic anemia.

1.4 Structural analysis of the spleen and liver

The increased erythropoiesis was confirmed by the marked presence of erythroid cells in spleen and liver histological sections of PKD mice (Figures 23 and 24), suggesting an intense extramedullar erythropoiesis in RPK deficient animals. We observed a predominance of erythroid cells in PKD spleen sections when compared to control groups, indicating an expansion of the red pulp in spleens from anemic mice (Figure 23). We also observed clusters of erythroid cells in the intersticial spaces outside the blood vessels of PKD livers, with a higher frequency in PKD than in C57BL/6 mice (Figure 24.A). Additionally, iron deposits were predominant in the liver of PKD animal, but not in C57BL/6 livers, indicating an acute iron overload in the anemic mice caused by the continuous hemolytic process (figure 25.B).

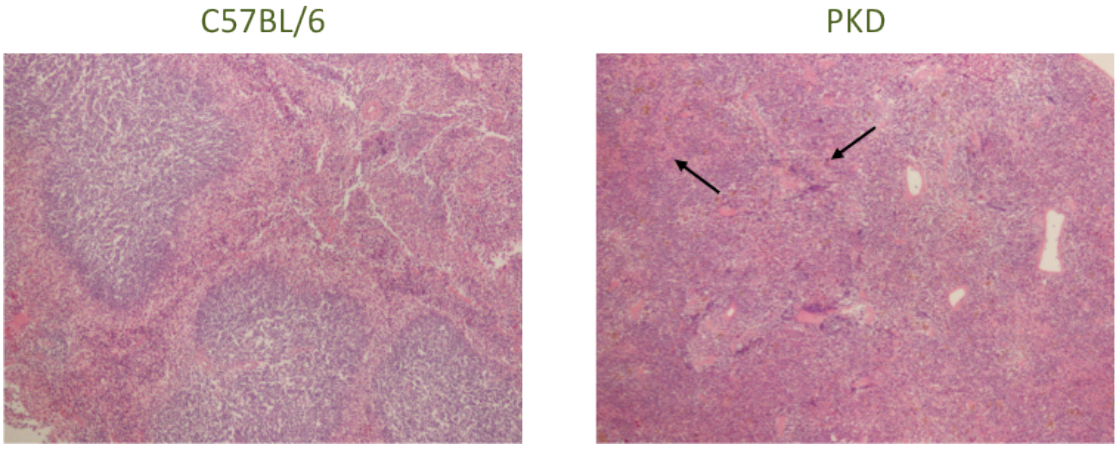

Figure 23: Study of spleen structure. Representative histology sections from C57BL/6 and PKD spleens stained with hematoxylin-eosin and photographed using 4x objective.

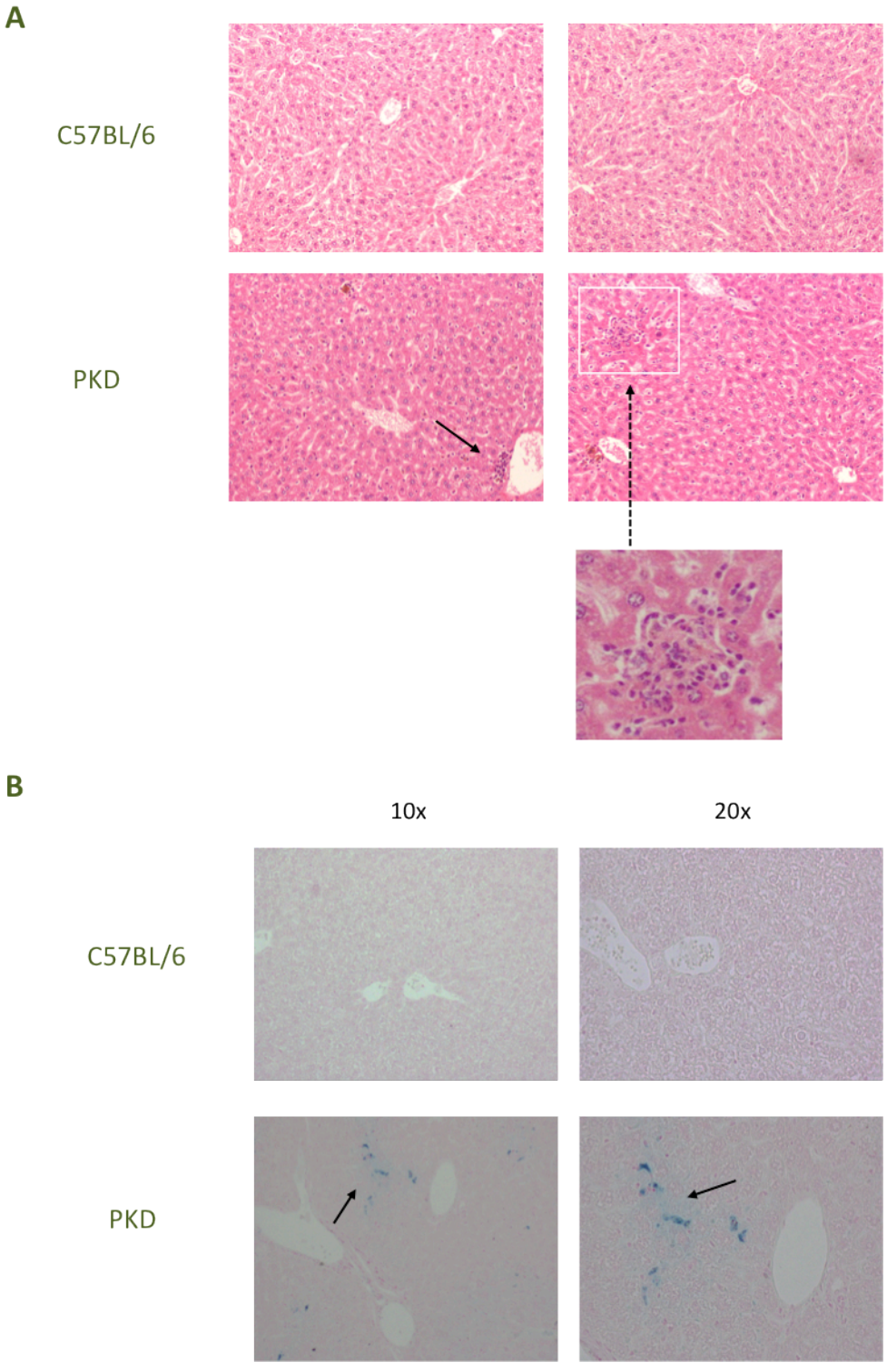

Figure 24: Histological analysis of the liver. Livers of C57BL/6 and PKD mice were harvested and analyzed by staining with **A)** hematoxylin-eosin to highlight the tissue structure or **B)** with Prussian Blue staining to detect iron deposits. Photographs were taken using 10x or 20x objectives in a light microscope. The arrows point to erythroid cells clusters indicative of extramedullar erythropoiesis in figure A or iron deposits in figure B.

1.5 Study of the hematopoietic progenitor compartment

Hematopoietic progenitor compartments from bone marrow and spleen were studied in the different mouse strains to determine if the anemic context was affecting the hematopoiesis in earlier stages of the differentiation process. First, we studied the hematopoietic progenitor content by CFU assays seeding either total BM or spleen cells on methylcellulose-based medium, and scoring the number of CFUs that had grown after 7 days in culture. The medium used contained rhEpo and was formulated to support optimal growth of erythroid progenitors (BFU-E), granulocyte-macrophage progenitors (CFU-GM, CFU-G and CFU-M) and multipotential granulocyte, macrophage, erythroid and megakaryocyte progenitors (CFU-GMM) colonies. Because we did not observe clear hemoglobinization of the erythroid colonies in these cultures, we recorded the number of total CFUs, including colonies committed to different lineages (Figure 25.A). RPK deficient animals showed a significant reduction in the number of bone marrow CFUs comparing to C57BL/6 or AB6F1 mice, but in the spleen the situation was the opposite, showing a significantly higher number of CFUs. This trend was also observed in comparisons with A/J mice, but the data were not significant. We also noticed that PKD CFU colonies from either BM or spleen were smaller than those from control groups (data not shown).

To study whether the anemic condition and compensatory erythropoiesis in PKD mice could affect BM progenitors behaviour, we performed hematopoietic progenitor mobilization protocols. Control C57BL/6 mice and anemic PKD animals were injected with PBS or PEG GM-CSF in three independent experiments (Figure 25.B), and the number of CFUs was recorded in peripheral blood 4 days after the stimulation. We observed that deficient PKD mice showed higher numbers of progenitors in PB than C67BL/6 mice, although it was not statistically significant. However, mobilization capacity was significantly reduced in PKD mice after stimulating the bone marrow with PEG G-CSF when compared to mobilized control mice, demonstrating an impairment of the BM mobilization mechanisms in RPK deficient mice, in spite of having an increased basal mobilization. Whether this finding is caused by the hemolytic anemia status or by the RPK deficiency *per se*, will require further experiments.

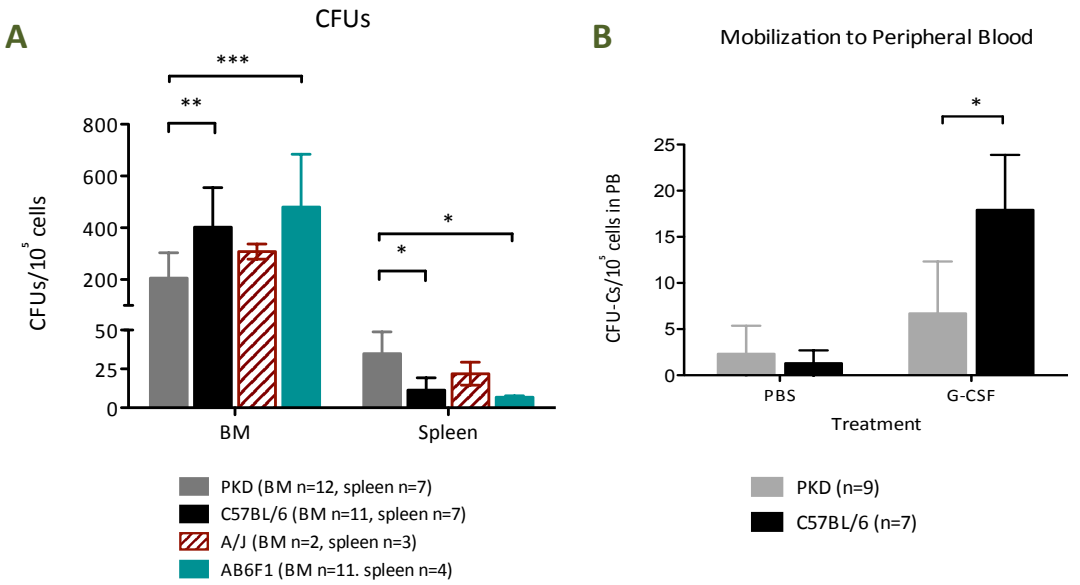

Figure 25: Hematopoietic progenitors studies. **A)** Recorded CFUs from BM and spleen hematopoietic progenitors cultures in PKD, C57BL/6, A/J and AB6F1 mice. Data represent the average and SD in each group, statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$). **B)** Recorded CFUs in peripheral blood from C57BL/6 and PKD animals four days after the injection of PBS vehicle or PEG G-CSF mobilization agent. Data represent the average and SEM in each group, statistically analyzed by non-parametric Mann-Whitney test ($p<0.05$).

To study whether RPK deficiency could impair the undifferentiated progenitor compartment, we studied the LSK ($Lin^-Sca1^+cKit^+$) and Slam ($Lin^-CD150^+CD48^-$) populations in bone marrow and spleen from the different animals (Figure 26.A). In both hematopoietic organs, LSK percentage was lower in PKD animals when compered to controls, although the differences were only statistically significant in the BM. Unexpectedly, the lowest percentage of spleen LSK progenitors was found in A/J mice. Slam progenitors were also reduced in the BM, with a percentage in spleen equivalent to that of C57BL/6 mice (Figure 27.B). Despite the reduction of uncommitted progenitors observed in the BM of PKD animals, the cellularity of this organ did not change among groups (Figure 27.C). On the contrary, a significant increase of spleen cellularity was found in RPK deficient animals compared to controls values.

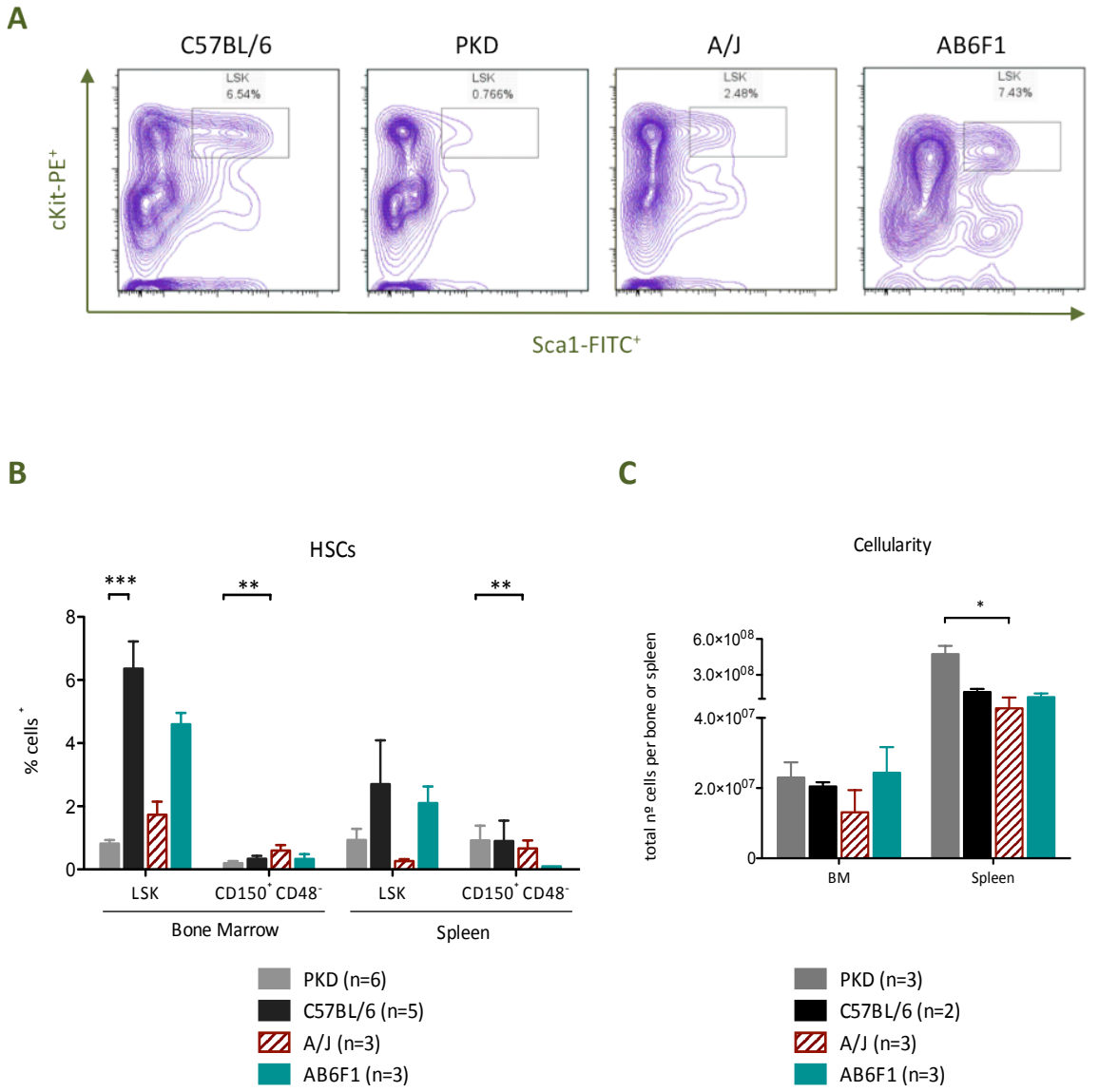

Figure 26: Study of uncommitted hematopoietic progenitors. Flow cytometry study of LSK and Slam progenitors from bone marrow and spleen following two different strategies. Sca1 and cKit expressing lineage negative progenitors in the upper part of figure, and CD150+ cells within the lineage negative and CD48- population in the lowest part. **A)** Representative flow cytometry histograms from BM LSK progenitors. **B)** Quantification of bone marrow and spleen progenitors from PKD, C57BL/6, A/J and AB6F1 mice studied following two different flow cytometry strategies described above. **C)** BM and spleen cellularity of PKD, C57BL/6, A/J and AB6F1 mice. Average and the SEM in each group are represented in both graphs. Data were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

2. Gene therapy assays in PKD mouse model

Once we confirmed that AcB55 congenic mice reproduced the hematological symptoms of human PKD, we performed different gene transfer experiments to develop an efficient and safe pre-clinical gene therapy for RPK deficiency. We previously demonstrated the feasibility of gene therapy as a therapeutic option for this deficiency in the AcB55 mouse model using γ-RV vectors (Meza et al., 2009). However, lentiviral vectors have been proven to have a safer integration pattern and therefore, a lower oncogenic potential than γ-RV vectors. In addition, like in other inherited metabolic diseases, the main limitation in developing a gene therapy protocol for PKD is the frequent lack of selective advantage of genetically corrected cells. This implies that a high number of corrected cells and a high level of transgene expression are required to achieve a therapeutic benefit. In order to overcome these limitations, we assayed two different gene therapy approaches based on HIV-derived lentiviral vectors, which were engineered carrying two different ubiquitous promoters (hPGK or CMV) and the human RPK transgene, achieving different levels of phenotype correction.

2.1 Development of gene therapy protocols for PKD with self-inactivating lentiviral vectors carrying the CMV viral promoter

As a first approach, we constructed a HIV-based vector to express the full-length human RPK cDNA under the control of the CMV strong viral promoter (Figure 11). As a control vector we used a previously validated self-inactivating pRRL lentiviral backbone, which carried the EGFP transgene and the Wpre sequence to increase the stability of the mRNA (VandenDriessche et al., 2002). Vectors were pseudotyped with VSV-G envelope and concentrated by ultracentrifugation, obtaining titers ranging 2.10^7 to 10^8 vp/mL. Figure 27 shows that developed vector was functional in non-expressing mouse 3T3 cells, since human RPK protein was detected in pRRL-CMV-hRPK transduced cells but not in cells transduced with the EGFP control vector.

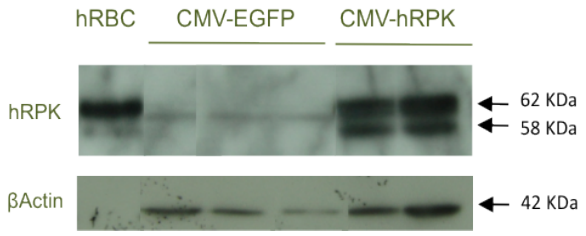

Figure 27: Western blot analysis of 3T3 cells infected with lentiviral vectors. Total protein extracts were obtained from 3T3 cells 8 days after the infection with pRRL-CMV-EGFP or pRRL-CMV-hRPK vectors. Samples were analyzed by Western blotting for human RPK protein and for mouse β-Actin as loading control.

To assess the efficacy of the lentiviral vector developed, Lin^- cells from BM of PKD male mice were harvested and transduced with EGFP or hRPK expressing vectors, and were next transplanted into lethally irradiated deficient female mice. Erythroid variables (RBC counts, HTC index, HGB, MCV, HCV, Epo levels and reticulocyte levels) and splenomegaly, as well as other molecular parameters (VCN, % chimerism, ectopic hRPK expression) were measured throughout the time and compared with wild type and non-transplanted PKD littermates. For optimizing the gene transfer protocol we performed a total of 5 experiments following different conditions that are showed in table 8.

In experiment number 1, no pre-stimulation of donor hematopoietic progenitors was performed and Lin^- cells were transduced in a single infection of MOI 3 (Table 8). With these conditions, PKD mice transplanted with pRRL-CMV-hRPK transduced cells showed an increase of RBC counts and hematocrit index at 120 days post-transplantation (dpt), but these differences were not significant when compared to controls (Figure 28.A and B). Mice transplanted in parallel with pRRL-CMV-EGFP control vector showed an initial 20% transduction at 30 dpt, but decreased with time approaching zero at the end of the experiment (90 dpt) indicating a possible loss of donor engraftment (Figure 28.C).

EXPERIMENT	TRANSDUCTION			TRANSPLANTATION					COMMENTS
	Pre-stimulation	Infection cycles	MOI (pv/cell)	Cells Lin⁻ per mouse	Transduction % EGFP	VCN/cell WBC	Chimerism % SRY	hRPK Exp. relativ. mGAPDH	
1	No	1	3	$4.6x10^5$	20	n.d	n.d	n.d	Partial correction Loss of engraftment
2	Yes	2	2	$8.8x10^5$	0.8	n.d	n.d	n.d	No transduction No correction
3	Yes	3	10	$1.7x10^5$	80	2.03 ± 0.03	89.71 ± 1.52	0.09 ± 0.03	Correction RBC, EPO Retics, Splenomegaly
SECONDARY TRANSPLANTS				$1.5x10^6$ *	–	2.36 ± 0.66	30.56 ± 17.60	0.04 ± 0.02	Partial correction Loss of engraftment possible silencing
4	Yes	3	10	$4x10^4$	60	n.d	n.d	n.d	Partial correction Lost of engraftment
5	Yes	3	10	$9.2x10^5$	6	0.87 ± 0.05	90.54 ± 8.82	0.005 ± 0.00	No correction Silencing

n.d: non-determined * frozen total BM cells per mouse

Table 8: Summary of transduction and transplantation assays performed with lentiviral vectors carrying the CMV viral promoter and EGFP or human *PKLR* cDNA transgenes. Numbers represent the average and the standard error (SEM) in the different groups for the different parameters analyzed at 30 dpt. n.d non-determined; * total BM cells per mouse from frozen BM samples.

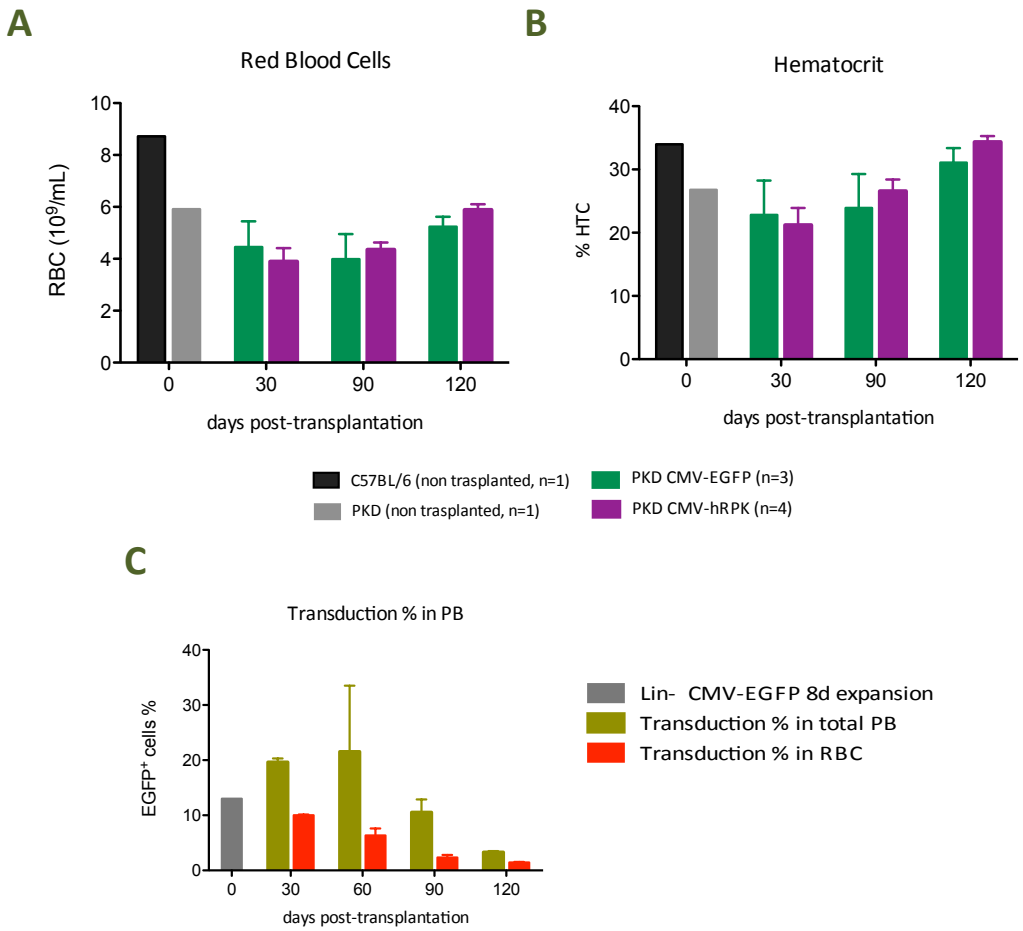

Figure 28: Analysis of hematological variables in animals transplanted with transduced cells in experiment 1. **A)** RBC counts and **B)** HTC percentage measured in transplanted mice along four months after transplant. **C)** Transduction percentage in peripheral blood of PKD mice transplanted with pRRL-CMV-EGFP transduced cells analyzed by flow cytometry at different time-points. Graphs represent the average and the standard error of the average (SEM). Data were not statistically analyzed because sample size was < 3 in non-transplanted control groups.

With the aim of improving the transduction efficacy, in experiment 2 we pre-stimulated the donor Lin^- cells with mSCF and hIL11, but no efficient transduction was achieved (Table 8). In experiment 3 we increased the number of infections up to three using a MOI of 10 and achieving phenotype correction as soon as 30 dpt. In the control group, an initial 80% EGFP transduction was observed in CFU progenitors, which stayed constant throughout the experiment in both PB and CFUs from BM and spleen of control mice transplanted with pRRL-CMV-EGFP transduced cells (data not shown).

Donor chimerism was determined in blood cells from all transplanted mice using a calibrated curve of male/female WBC as previously described by Navarro *et al* (Navarro et al., 2006), and was estimated on 90% donor engraftment in both groups. The average number of vector copies per cell (VCN) was also calculated, detecting 2.35 ± 0.03 copies per cell in BM and 2.03 ± 0.01 copies per cell in PB of mice transplanted with pRRL-CMV-hRPK transduced cells (Table 8). Thus, with 2 vector integrations per cell and 90% of blood donor cells we observed an improvement of the anemic phenotype in mice transplanted with genetically corrected cells. As soon as 30 dpt the hematological variables in these mice reached values close to C57BL/6 wild-type controls, showing an increase of RBC counts, HGB and HTC, and a decrease in MCV and MCH variables Figure 29.A-E), but no changes were found in WBC counts comparing with PKD anemic controls (data not shown). On the contrary, the abnormal red cell phenotype was not corrected in animals transplanted with cells transduced with the EGFP control vector, which behaved like the PKD anemic controls. Additionally, we observed a marked improvement of reticulocyte percentage in mice transplanted with pRRL-CMV-hRPK transduced cells, which remained stable (Figure 30).

Splenomegaly, another typical sign of PKD disease, was also corrected. We observed that PKD animals transplanted with pRRL-CMV-hRPK transduced cells showed a marked reduction in spleen weight and size at 90 dpt (Figure 31.A and B), accompanied by a decrease in plasma and serum Epo levels (Figure 31.C). These results indicated a reduction of the hemolytic process in PKD transplanted mice, and therefore the rescue of the PKD anemic phenotype, as a consequence of the coRPK transgene expression (0.09 ± 0.03 hRPK expression relative to mGAPDH) in PB cells.

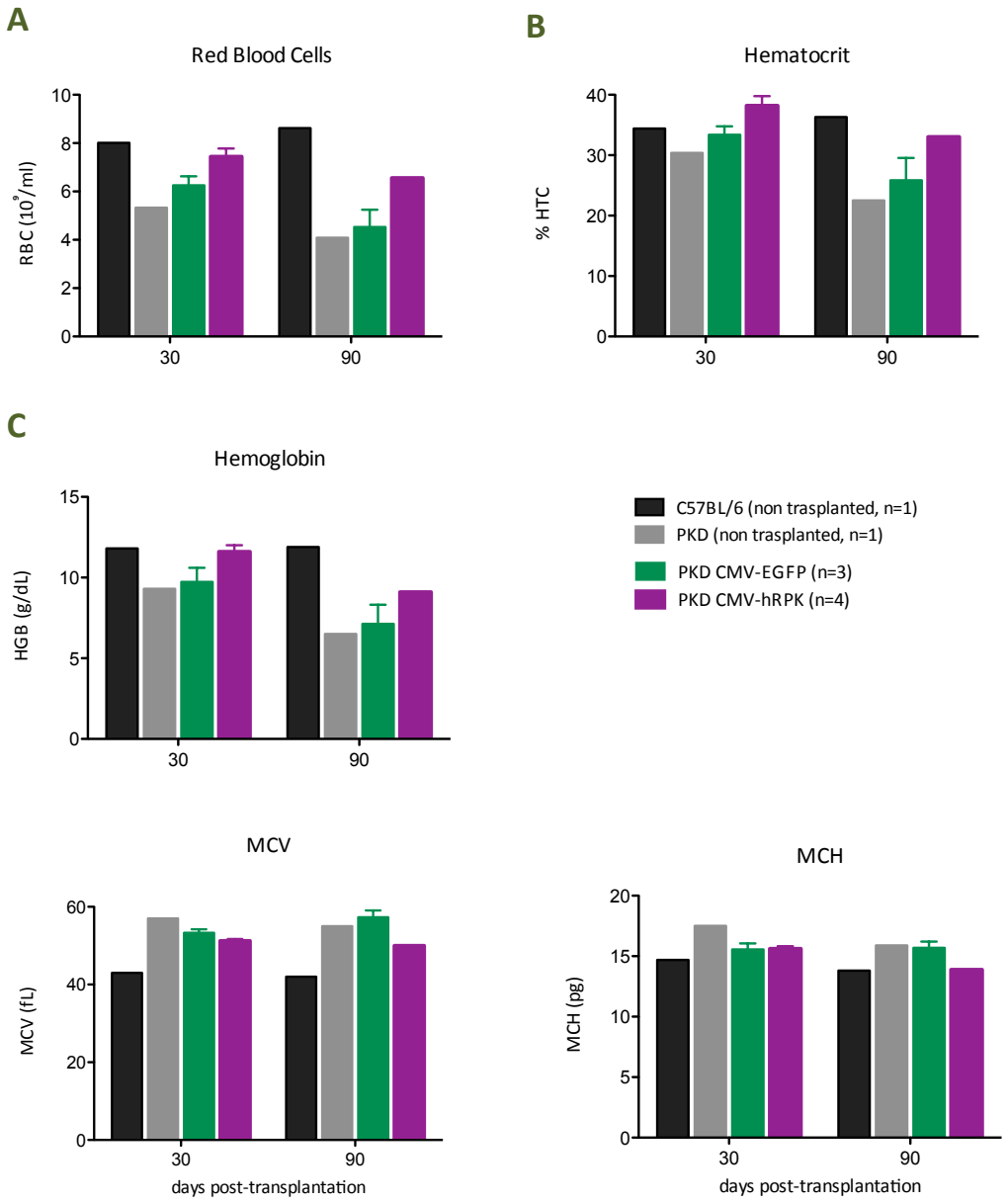

Figure 29: Analysis of hematological variables in animals transplanted with transduced cells in the experiment 3. **A)** RBC counts, **B)** HTC percentage, **C)** HGB levels, **D)** MCV and **E)** MCH indexes measured in peripheral blood of PKD transplanted mice with either pRRL-CMV-EGFP or pRRL-CMV-hRPK lentiviral vector at different months after the transplant. Control groups C57BL/6 and PKD littermates were analyzed in parallel. Bars correspond to the average and SEM per group. Data were not statistically analyzed because n < 3 in non-transplanted control group.

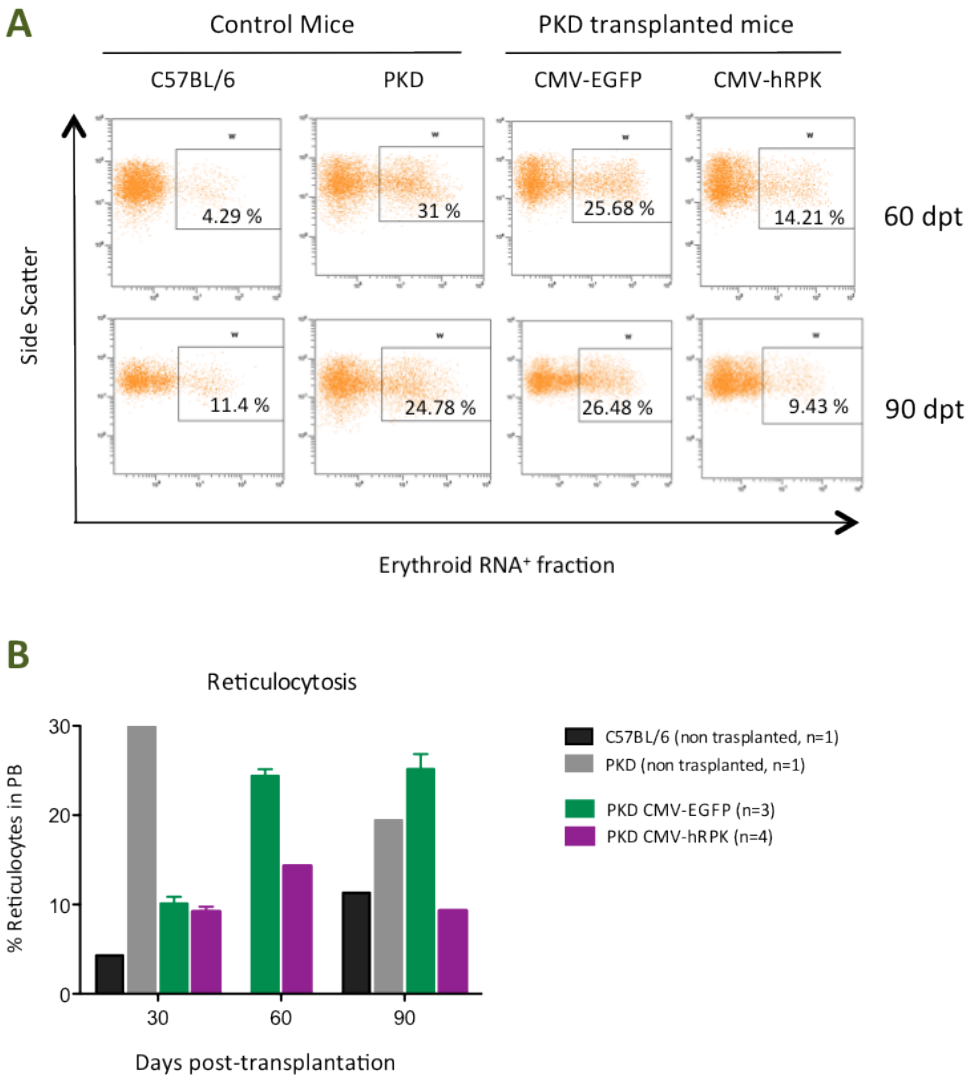

Figure 30: Study of reticulocyte levels of PKD transplanted mice from the experiment number 3. **A)** Representative dot-plot of the flow cytometry analysis used to monitor reticulocytosis in peripheral blood. **B)** Reticulocyte percentages recorded in the different groups throughout the experiment. Bars represent the average per group and SEM of the data. Data were not statistically analyzed because sample size was < 3 in non-transplanted control groups.

Despite observing an improvement in the hematological status of mice treated with the therapeutic lentiviral vector in the experiment number 3, only one mouse remained alive at the end-point due to causes other than experimental conditions. Therefore secondary receptors were transplanted individually with $1.5x10^6$ cells from primary receptors total BM, that were harvested and frozen at different time-points. Surviving secondary receptors (50%) showed a modest improvement of hematological parameters, including an increase of RBC counts and HGB levels (Figure 32.A and B) and a decrease of

MCV and MCH variables (Figure 32.E and D). A reduction of the reticulocyte percentage was also observed (Figure 32.C) in animals from the hRPK group, that maintained the number of proviral integrations in peripheral blood equal to primary transplanted mice (2.36 ± 0.66 copies per WBC).

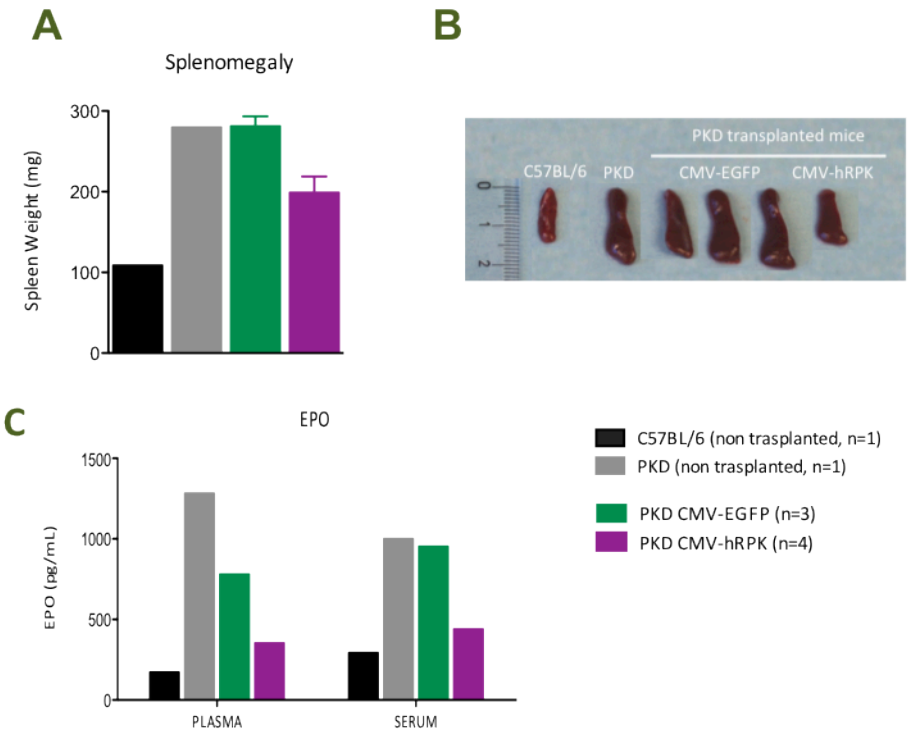

Figure 31: Study of spleen and circulating Epo levels in PKD mice transplanted from the experiment 3. **A)** Spleen weight of control mice and the two different transplanted PKD groups at 90 dpt. **B)** Pictures of spleens just after the extraction. **C)** Plasma and serum Epo levels in the different groups measured at 90 dpt by ELISA. Bars represent the average and the SEM of the average. Data were not statistically analyzed because sample size was < 3 in non-transplanted control groups.

Group	WBC (10^6/ml)	RBC (10^9/ml)	HTC %	HGB (g/dL)	MCV (fL)	MCH (pg)
C57 (n=1)	4.51	8.78	36.74	12.3	42	14
B19 (n=1)	2.85	3.34	18.65	5	56	15.1
EGFP (n=2)	1.75 ± 0.74	4.03 ± 1.48	24.67 ± 8.32	7.05 ± 2.47	61.5 ± 2.12	17.55 ± 0.21
hRPK (n=2)	1.51 ± 1.37	5.65 ± 0.52	30.97 ± 1.48	8.85 ± 0.35	55 ± 2.83	15.7 ± 0.85

Table 9: Hematological variables recorded in peripheral blood of transplanted and control animals from the experiment 4 at three months post-transplantation. Numbers represent the average and the standard error (SEM) in the different groups.

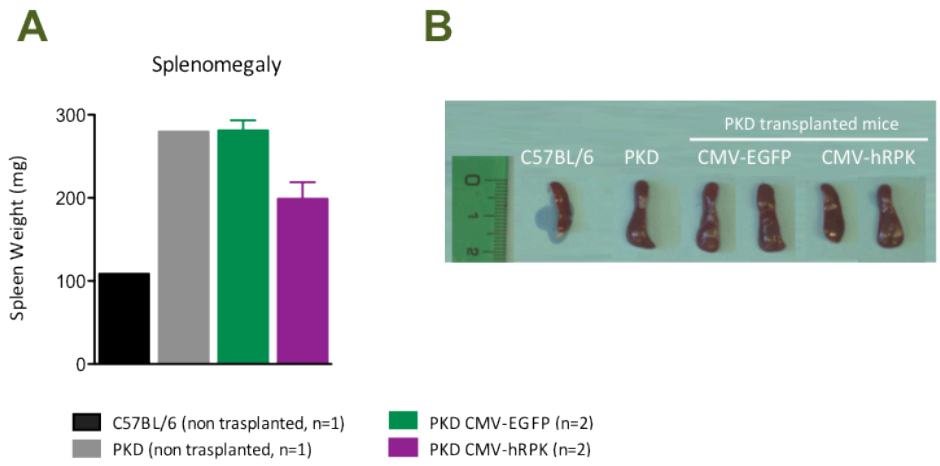

Figure 33: Study of spleen in transplanted and control mice of the fourth experiment at 90 dpt. Spleen weight **(A)** and photographs **(B)** of the different groups of animals. Bars represent the average of each group and the SEM of the data set. Data were not statistically analyzed because sample size was < 3.

The last CMV-based experiment (number 5) was also performed by pre-stimulating donor Lin^- cells and infecting in three cycles of MOI 10, although in this case infections were carried out every 12h to avoid cells differentiating and loosing their repopulating capacity (Table 8). Even so, only 8% of transduced Lin^- cells expressed the EGFP transgene after 8 days expansion, and this percentage decreased to 6% in mice transplanted with pRRL-CMV-EGFP vector at 30 dpt (data not shown). Transplanted mice from pRRL-CMV-hRPK group maintained the anemic phenotype the same way as EGFP control mice or PKD

non-transplanted mice, in spite of showing a 90% donor chimerism in PB. However, VCN in blood cells (0.87 ± 0.05 vector copies per cell) was half of vector integrations detected in experiment 3 (Table 8), and rendered 18-fold lower expression of the hRPK transgene. These findings suggest a possible silencing of the CMV promoter, confirming previous results using the same promoter (Oertel et al., 2003).

2.2 Development of gene therapy protocols for PKD with self-inactivating lentiviral vectors carrying the human PGK eukaryotic promoter

Our results showed that efficient and stable correction of the PKD disease was dependant on levels of donor chimerism and transgene expression. However, the silencing event observed and associated with the use of CMV promoter encouraged us to move forward to a safer and more physiological vector design. We generated a new SIN HIV-based lentiviral vector carrying the hPGK eukaryotic promoter (Figure 13), which drives a more stable and sustainable expression of the transgene than CMV promoter (Gerolami et al., 2000) (Gonzalez-Murillo et al., 2010). In addition, hPGK promoter was less genotoxic than viral promoters such as SFFV in studies of insertional transformation of hematopoietic cells in a lentiviral context (Modlich et al., 2009). This new vector design also included an optimized version of the human wild-type *PKLR* cDNA (coRPK) and a mutated version of the Wpre post-transcriptional element (Wpre*).

2.2.1 Correction of anemic phenotype in primary PKD transplanted mice

Developed lentiviral vectors carrying the EGFP or coRPK transgenes were pseudotyped with VSV-G and concentrated by ultracentrifugation. Lentiviral stocks of 10^7-4.10^9 vp/mL were used to develop two gene therapy assays, at low and high MOI respectively (MOI 1-10). We maintained the transduction protocol previously tested in the CMV assays achieving phenotypic correction. BM from RPK deficient male mice was harvested and isolated Lin^- cells were expanded for 24h in presence of rhIL11 and rmSCF (Figure 34). Two consecutive transductions of 24h were carried out with the pCCL-hPKG-EGFP-Wpre* control or pCCL-hPGK-coRPK-Wpre* therapeutic vector. PKD female receptor

mice were conditioned by lethal irradiation and 2.10^5 transduced Lin^- cells per mice were intravenously transplanted.

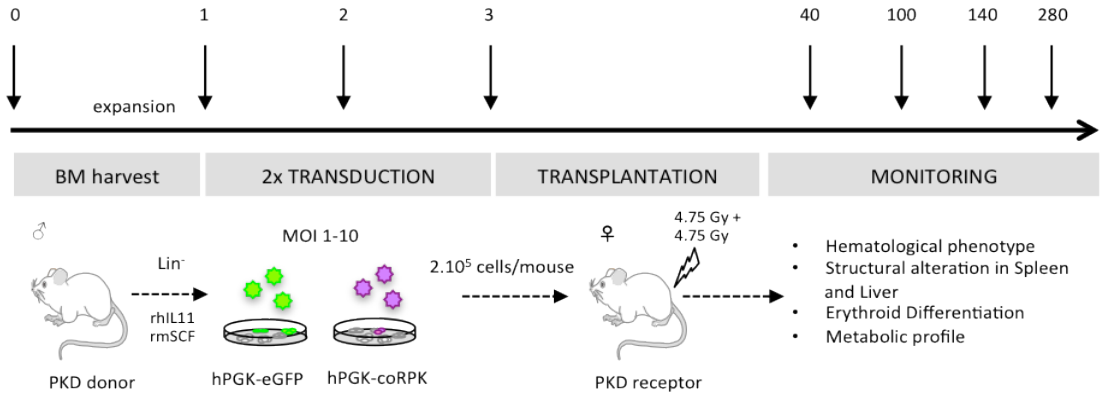

Figure 34: Scheme of the gene therapy assays performed with lentiviral vectors carrying the hPGK promoter. Numbers on the top represent days after BM harvesting. Correction of the phenotype was studied by the analysis of different hematological and metabolic parameter throughout 4 months after transplantation.

Transplanted animals were monitored for 4 to 9 months, achieving a stable and long-term phenotype correction in those expressing the coRPK transgene. These animals showed a significant improvement in RBC counts when compared with the EGFP control group as soon as 40 dpt, which was maintained for up to 9 months (Figure 35.A). Complete analysis of peripheral blood revealed the significant correction of the hematological phenotype in terms of HGB levels, HTC percentage and MCV and MCH indexes in PKD mice transplanted with genetically corrected cells, reaching values close to C57BL/6 healthy mice (Figure 35.B and Table 10). On the contrary, leucocyte levels (WBC) were not affected by the genetic modification. In parallel, transduction percentage was measured in control mice by flow cytometry analysis of EGFP expression, showing a stable transduction of 30-70% throughout the time in both experiments (data not shown).

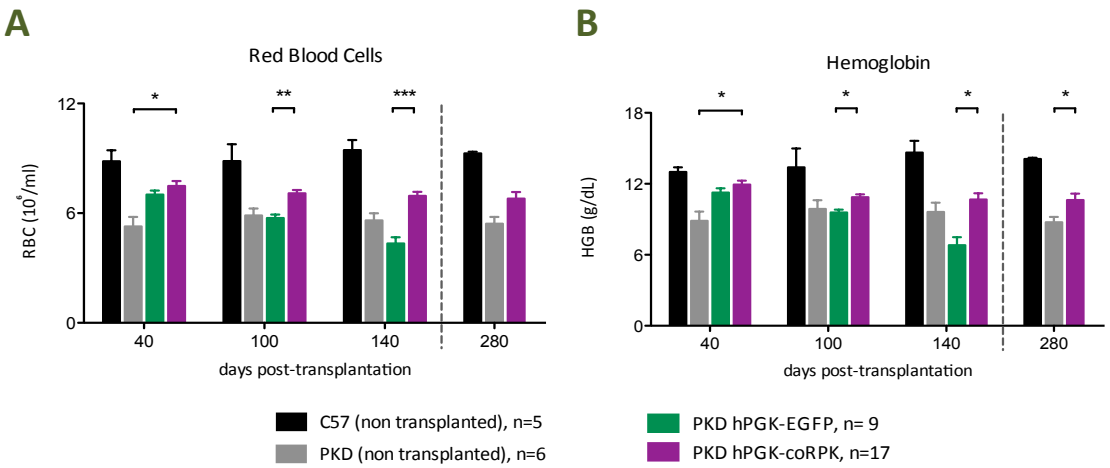

Figure 35: Erythrocyte and Hemoglobin levels in animals transplanted with transduced cells. **A)** RBC counts and **B)** HGB levels were measured in PB of PKD animals transplanted with cells previously transduced with pCCL-hPGK-EGFP or pCCL-hPGK-coRPK lentiviral vectors along 4 months post-transplantation. As controls, anemic PKD animals and C57BL/6 healthy mice were measured in parallel. A few animals were maintained for a further 5 months more (separated by a dashed line), demonstrating the long-term phenotype correction. Data from 280 dpt were not statistically analized because n < 3. Bars represent the average and SEM per group at the different time-points, and were analyzed by non-parametric Kruskal-Wallis statistical test ($p<0.05$). Data from 280 dpt which were analyzed by Mann-Whitney test ($p<0.05$).

Group	WBC (10^6/mL)	HTC (%)	VCM (fL)	HCM (pg)
C57BL/6 (n=5)	5.47 ± 0.58	36.15 ± 2.64	38.20 ± 0.86	15.44 ± 0.21
PKD (n=6)	3.54 ± 0.60	28.67 ± 2.21	51.17 ± 0.65	17.10 ± 0.42
PKD hPGK-EGFP (n=8)	2.14 ± 0.41	21.32 ± 1.51	49.13 ± 0.72	15.46 ± 0.62
PKD hPGK-coRPK (n=16)	3.02 ± 0.40	31.09 ± 1.45**	45.65 ± 0.84**	15.35 ± 0.53*

Table 10: Hematological variables recorded in peripheral blood of transplanted (PKD hPGK-EGFP or hPGK-coRPK) and non-transplanted animals (C57BL/6 and PKD mice). Numbers represent the average and the SEM in the different groups at 140 days post-transplantation, and were analyzed by non-parametric Kruskal-Wallis statistical test ($p<0.05$). Differences in HTC percentage were significant when compared to PKD hPGK-EGFP mice, whereas VCM and HCM were significant when comparing with PKD anemic mice.

Reticulocytosis levels were also corrected in anemic mice expressing the coRPK transgene. These animals showed a significant decrease of circulating reticulocytes (Figure 36) and approached the levels of C57BL/6 control mice, while PKD animals transplanted with the control vector still showed marked reticulocytosis in parallel with PKD mice.

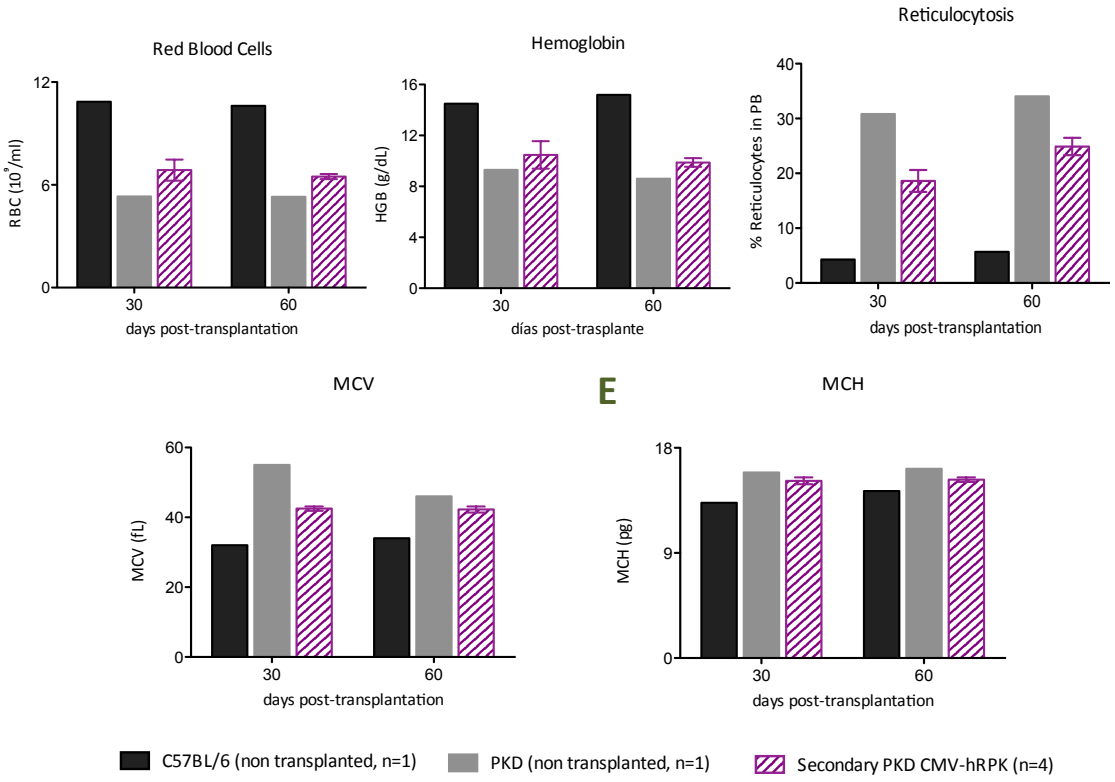

Figure 32: Analysis of hematological variables scored in peripheral blood of secondary receptor mice from experiment 3. PKD mice were transplanted after lethal irradiation with total BM cells from transplanted animals of the third experiment belonging to CMV-hRPK group. **A)** RBC counts, **B)** HGB levels, **C)** reticulocyte percentages, **D)** MCV and **E)** MCH indexes were measured in peripheral blood along two months post-transplantations. Bars represent the average and the SEM of the data. Data were not statistically analyzed because sample size was < 3 in non-transplanted control groups.

With these results, we performed a new experiment (number 4) maintaining the transduction conditions from the previous one, but again we achieved no significant correction of the PKD symptoms. In control mice, EGFP expression was monitored, showing a decrease of the transduction percentage with the time. PKD mice transplanted with pRRL-CMV-hRPK transduced cells showed an improvement of all hematological parameters apart from reticulocyte levels, but differences were not significant when comparing with the EGFP group (Table 9), probably due to either low levels of donor engraftment detected or poor transduction. Nonetheless, the spleen weight and size dramatically decreased in the therapeutic group approaching C57BL/6 values (Figure 33).

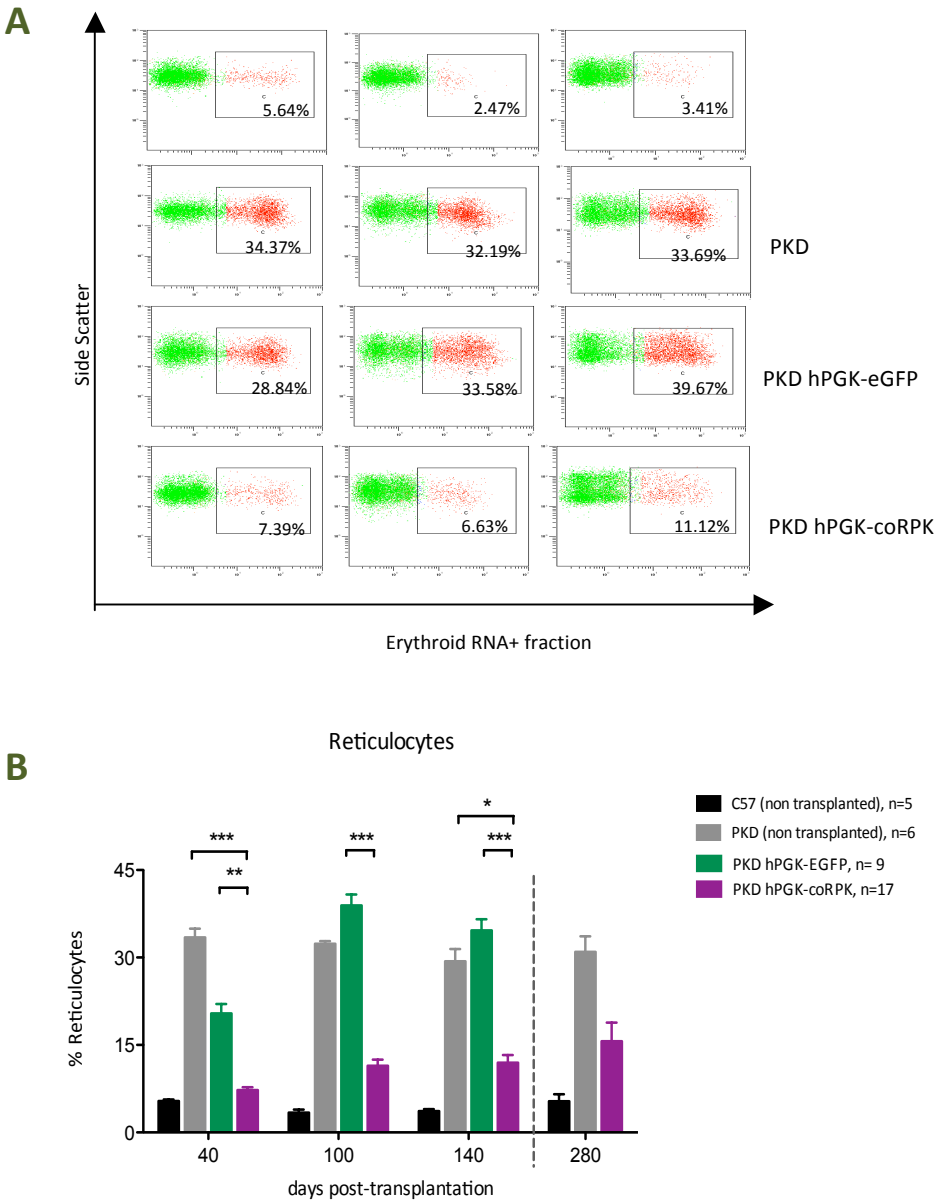

Figure 36: Study of reticulocyte levels in peripheral blood of the different mice analyzed. **A)** Representative dot-plot of the flow cytometry analysis used to identify the reticulocyte population by Acridine Orange staining. **B)** Reticulocyte percentages scored at different days post-transplantation in non-transplanted and transplanted mice. Few animals were maintained up to 280 dpt, demonstrating the long-term correction of reticulocytosis. Data from 280 dpt were not statistically analyzed because n < 3. Bars represent the average and SEM per group at different time-points. Data were analyzed by non-parametric Kruskal-Wallis statistical test ($p<0.05$), except at 280 dpt which were analyzed by Mann-Whitney test ($p<0.05$)

Additionally, biotinylated RBC lifespan also improved with the gene therapy treatment, being significantly extended in genetically corrected mice compared to non-transplanted PKD animals during the first 12 days after *in vivo* RBC labelling (Figure 37). In

both non-transplanted and transplanted PKD mice the initial percentage of biotinylated RBC was notably lower than in C57BL/6 mice (Figure 38.A), but PKD hPGK-coRPK animals showed a marked increment of the initial Biotin+ RBC as soon as 2 days after labelling (Figure 38.B), and their biotinylated RBC half-life was extended to 25 days. Unexpectedly, the half-life of labelled RBC was similar in anemic and healthy controls (19 days).

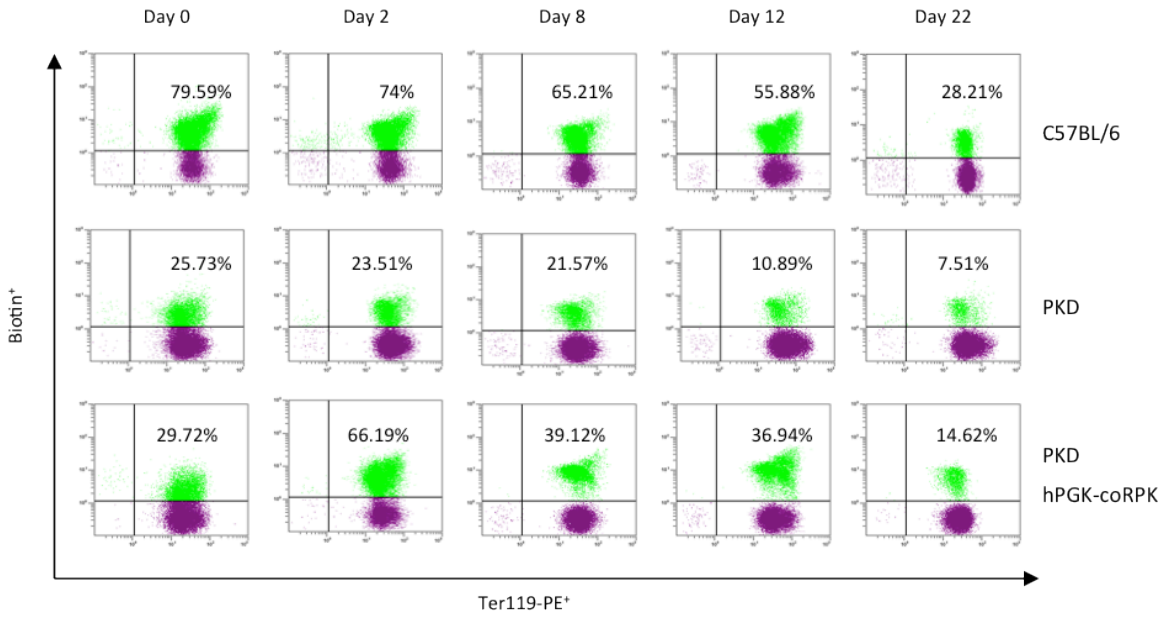

Figure 37: Study of red blood cells survival. At 140 days post-transplantation, control and transplanted mice were injected with biotin 3-sulfo-N-hydroxysuccinimide ester sodium salt and peripheral blood was periodically analyzed for 40 days. Figure shows the representative dot-plot of the flow cytometry analysis used to identify the biotin+ RBC within the total RBC population.

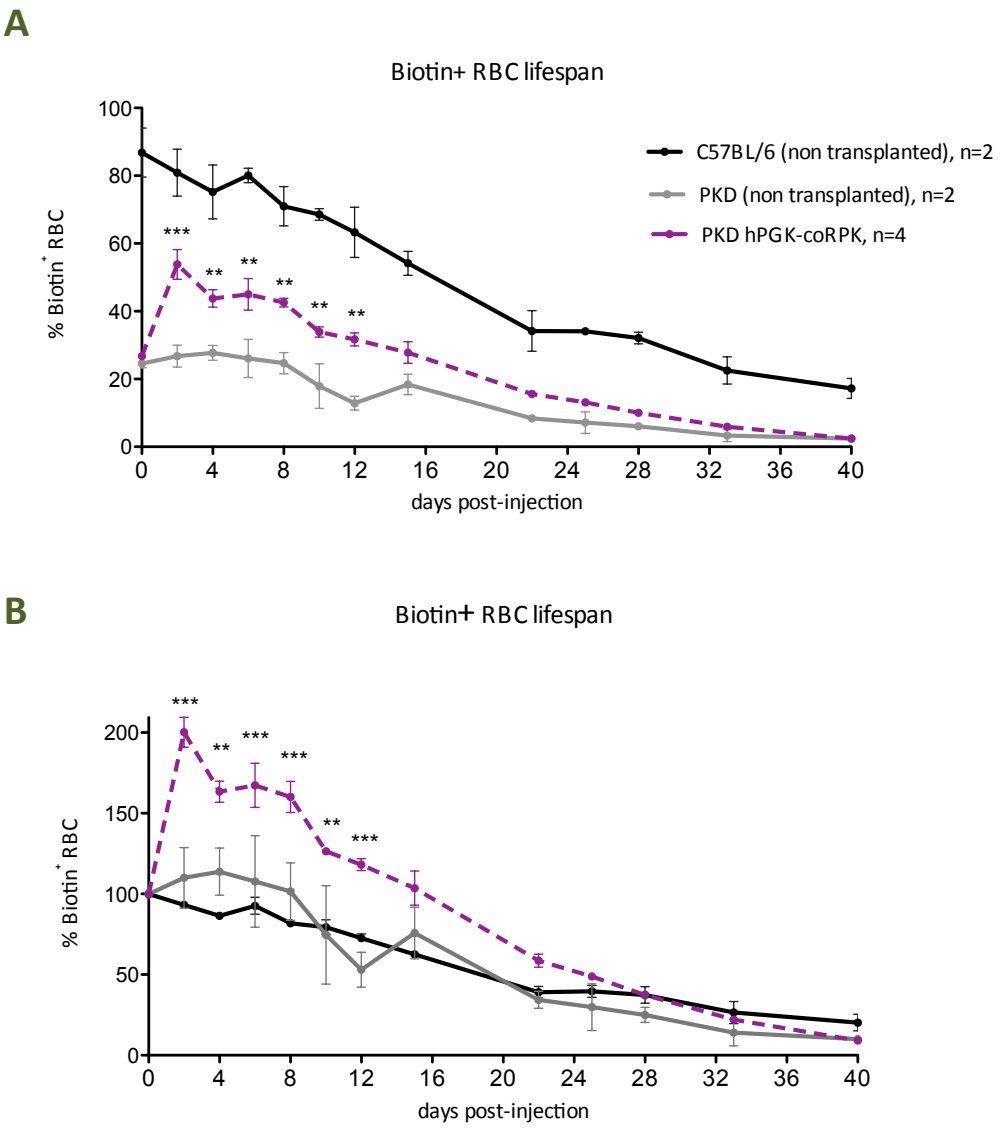

Figure 38: Kinetic of red blood cell survival. **A)** Direct data of biotin+ RBC percentages within the total RBC population over the time course. **B)** Relative biotin+ RBC survival within the total RBC population. The percentage obtained at day 0 was considered as 100% for each individual mouse and values measured along the time were normalized to day 0. Data represent the average and SEM in each group and were analyzed by two-way ANOVA statistic test ($p<0.005$)

2.2.2 Correction of splenomegaly and structural alterations of the spleen and liver in primary PKD transplanted mice

Splenomegaly was also reverted in deficient mice transplanted with genetically corrected cells, reaching spleen weights and sizes comparable to those of C57BL/6 controls (Figure 39). We also observed erythroid cell clusters in spleen sections from coRPK expressing mice with a low frequency than in deficient or EGFP transplanted PKD mice (Figure 40). This may be related to a decreased hemolytic process in PKD animals carrying the ectopic RPK protein, according to the assumed extended lifespan of corrected RBC in this model. On the other hand, the study of liver sections revealed a reduction of RBC accumulations in animals transplanted with cells transduced with the therapeutic lentiviral vector, as well as a reduction in the iron overload (Figure 41). Overall these data point to a reduction in the hemolysis process and in extramedullar erythropoiesis as direct consequences of the RPK transgenic expression.

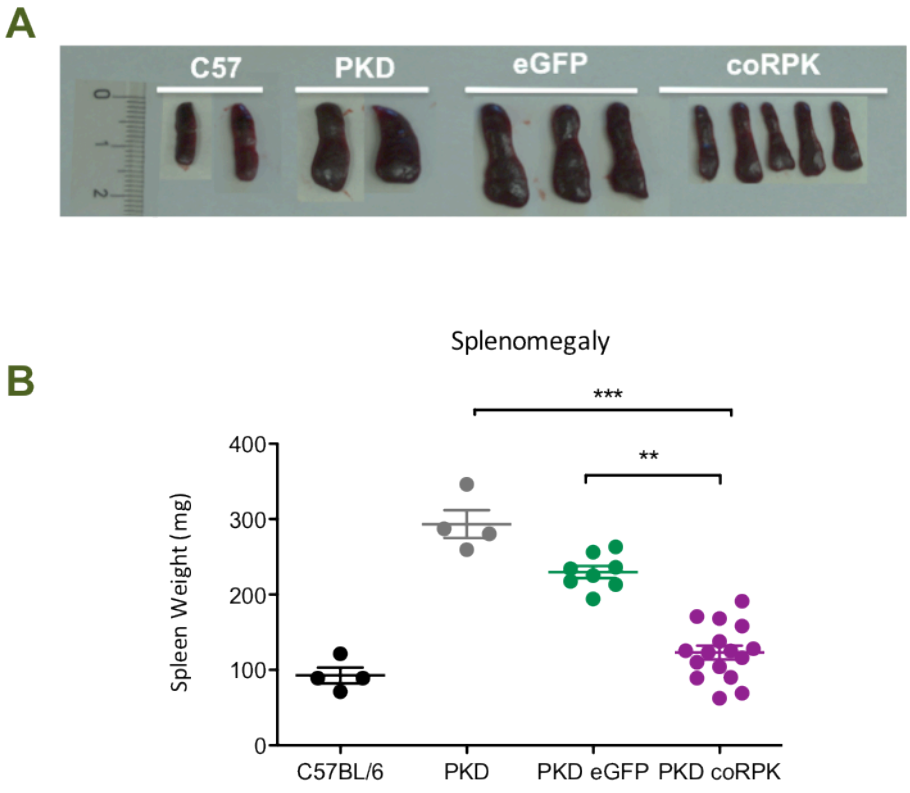

Figure 39: Study of spleen from transplanted and non-transplanted mice. Spleens from different groups of animals were analyzed at 140 dpt. **A)** Pictures of representative spleens. **B)** Spleen weight from the four different groups of mice. Data represent the individual values, the average and SEM in each group. Data were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

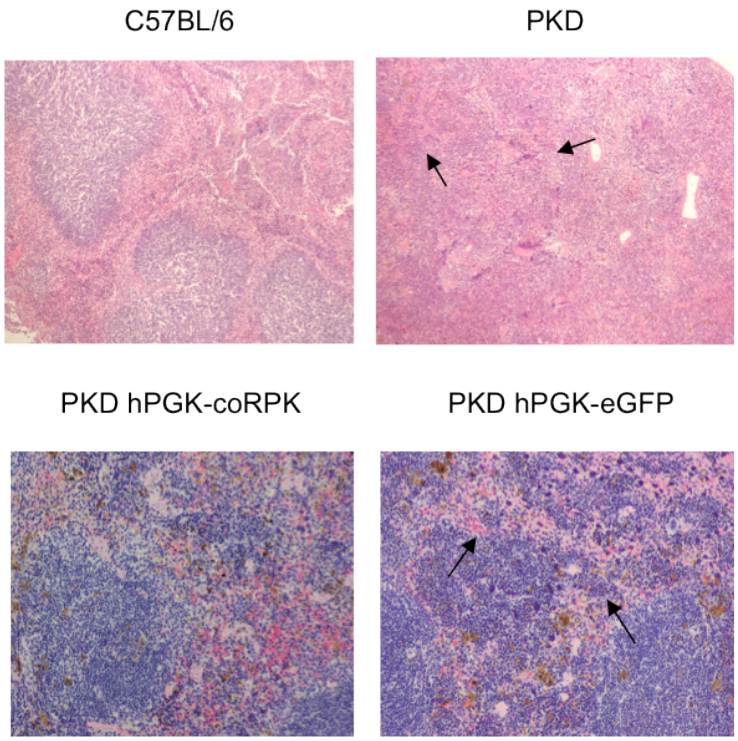

Figure 40: Study of spleen structure. Representative histology sections of spleens from the different groups of mice stained with hematoxylin-eosin and photographed using a 4x objective. Arrows indicate erythroid cell clusters indicative of extramedullar erythropoiesis.

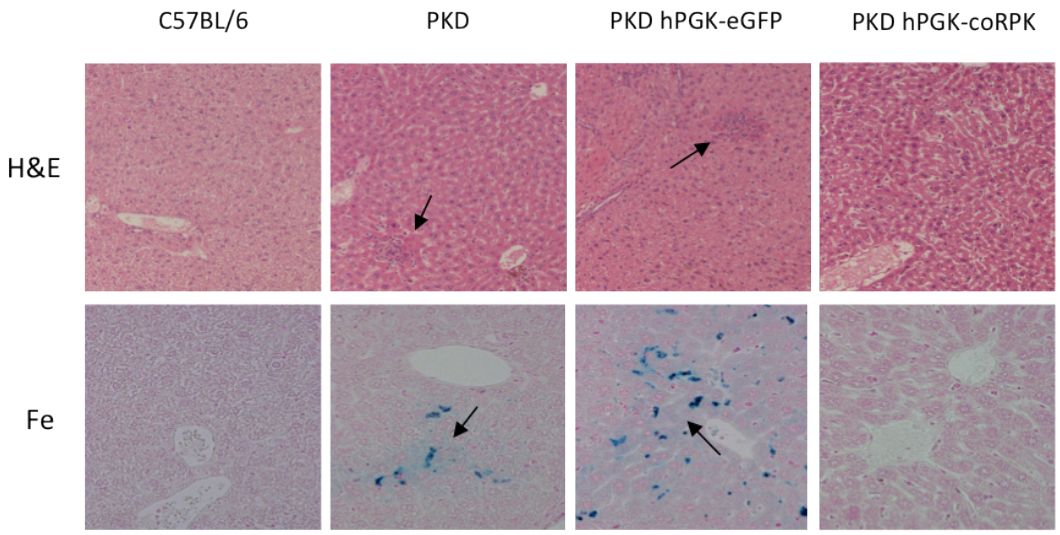

Figure 41: Histological analysis of the liver. Livers from non-transplanted C57BL/6 and PKD mice or genetically modified PKD mice were analyzed at 140 dpt by staining with hematoxylin-eosin staining (H&E) or with Prussian Blue staining (Fe) to detect iron deposits. Photographs were taken using 20x objectives in a light microscope. Arrows point out erythroid cells clusters (upper part) or iron deposits (lower part) indicative of extramedullar erythropoiesis and reduction of hemolysis respectively.

2.2.3 Correction of erythrocyte maturation in primary PKD transplanted mice and CFUs progenitors contents

To determine whether the correct expression of RPK in erythroid cells was able to revert the compensatory erythropoiesis process observed in PKD animals due to their continuous hemolytic state (Figure 40), we studied the differentiation patterns of the erythroid compartment in control and transplanted PKD mice by flow cytometry (figure 42). Deficient animals transplanted with coRPK expressing cells showed a reduction in immature erythroid precursors of proerythroblasts and basophilic erythroblasts (I and II subpopulations) in BM (Figure 43.A) and spleen (Figure 43.B), in parallel with a significant increase of the latest erythroid compartment (IV) equivalent to C57BL/6 control mice. Thus, the genetic modification with the vector carrying the RPK protein, stabilized the erythroid maturation in hematopoietic organs to normal levels, but not in those animals transplanted with control vector-transduced cells, that maintained the stressed erythropoiesis as a mechanism to compensate the anemia.

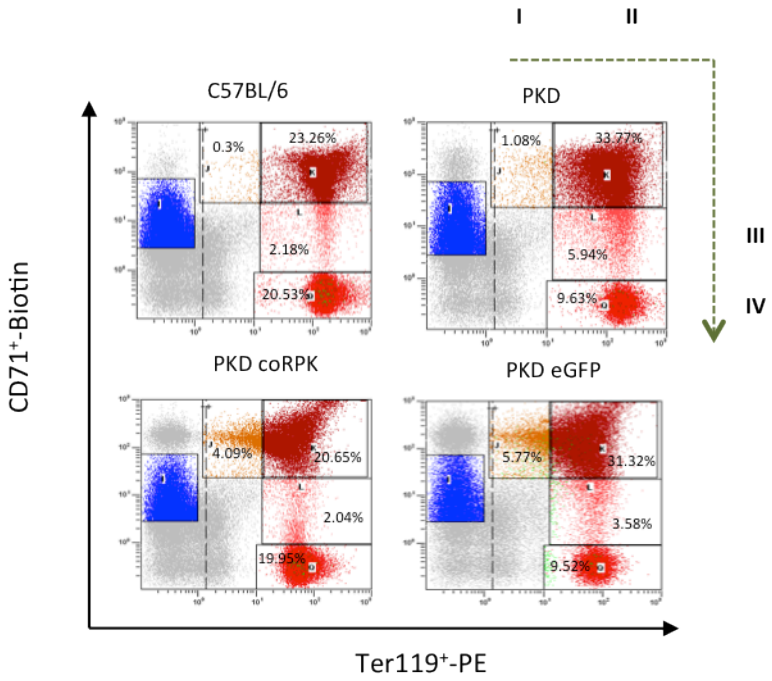

Figure 42: Flow cytometry study of the different erythroid subpopulations in transplanted and non-transplanted mice. Flow cytometry analysis of the expression intensities of Ter119 and CD71 markers distinguishes four erythroid subpopulations. I: Early proerythroblasts (Ter119med CD71high), II: basophilic erythroblasts (Ter119high CD71 high), III) late basophilic and polychromatophilic erythroblasts (Ter119high CD71 med) and IV: orthochromatophilic erythroblasts, reticulocytes and mature erythroid cells (Ter119high CD71low). Figure shows representative dot plots from the four groups at 140 dpt.

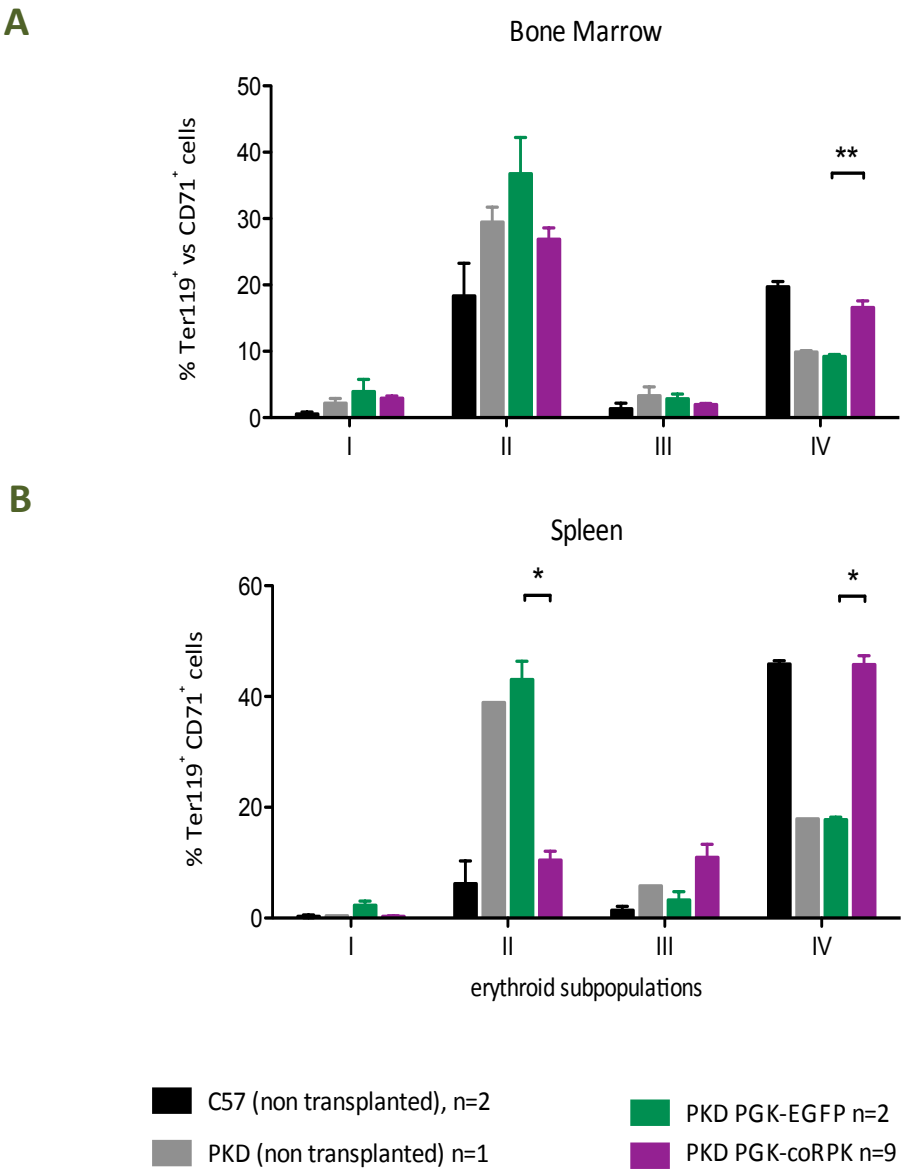

Figure 43: Study of erythroid differentiation in control and transplanted mice. Percentages of the different **e**rythroid subpopulation in **A)** bone marrow and **B)** spleen of non-transplanted mice and animal transplanted with control or therapeutic vector-modified cells. Data represent the average and the SEM in each group, and were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

The normalization of erythropoiesis in deficient animals treated with the therapeutic vector was accompanied by a reduction in the spleen CFUs (Figure 44.B) similar to C57BL/6 control animals. However, the expression of coRPK transgene did not revert the low content of progenitors found in the BM (Figure 44.A), showing equal levels to deficient animals or PKD mice carrying the control transgene.

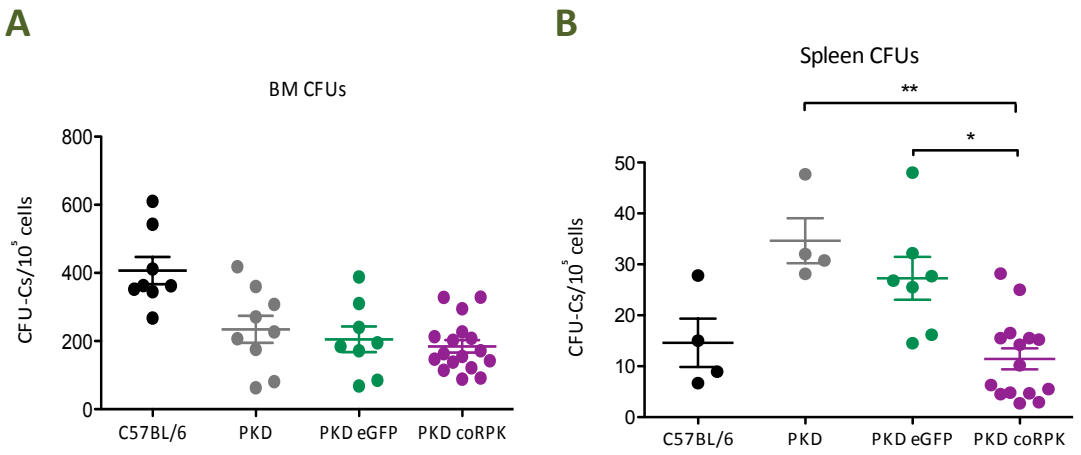

Figure 44: Hematopoietic progenitor assays in control and transplanted animals. **A)** Bone marrow and **B)** spleen CFUs at 140 dpt. Data represent individual values, average and SEM in each group, and were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

2.2.4 Correction of Epo levels in primary PKD transplanted mice

Additionally, a significant reduction of Epo levels was found in the plasma of PKD animals transplanted with coRPK expressing cells (Figure 45). As expected, PKD mice carrying the control transgene still maintained high Epo levels similar to deficient mice. These results demonstrated that genetically corrected animals suffer less acute hemolysis in PB and therefore, the accelerated RBC replenishment by compensatory erythropoiesis, partially controlled by Epo, was not needed in mice expressing the ectopic RPK protein.

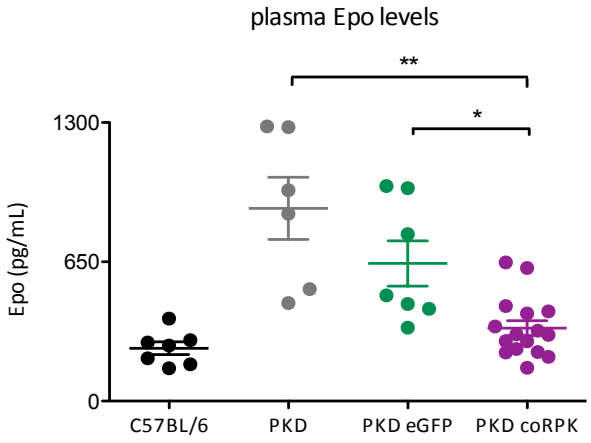

Figure 45: Plasma Epo levels in the different cohorts of mice at 140 dpt. Data represent individual values of Epo concentration measured by ELISA, the average and SEM in each group, and were statistically analyzed by non-parametric Kruskal-Wallis test ($p<0.05$).

2.2.5 Molecular parameters of transplanted mice

Through the genetic modification of HSC with the hPGK-coRPK lentiviral vector we addressed the functional reversion of the PKD hematological symptoms in two independent experiments. In assay number 2, stable and long-term phenotypic correction was achieved with a donor chimerism estimated on 60%, and this level was comparable in both experimental groups (Table 11). Assay number 1 also yielded a high reversion of the PKD phenotype, with 90% of CFUs progenitors carrying the provirus in the therapeutic group. However, we observed that the level of phenotypic correction was higher in experiment 1, consistent with the higher number of integrations per cells (4.76 ± 0.28 provirus copies/cell) and higher percentage of corrected CFUs in BM (91.08%) estimated in transplanted mice from experiment 2.

2.2.6 Secondary transplants

Once we demonstrated the efficacy of this pre-clinical protocol for the PKD, we explored its stability and safety by performing secondary transplants (Table 12). Although lentiviral vectors have seen to have lower oncogenic potential than conventional retroviral vectors (Montini et al., 2009); (Modlich et al., 2009), insertional mutagenesis is still a potential risk when using integrative vectors. 120-140 days after primary transplantation, $5x10^5$ to $4x10^6$ pooled total BM cells from primary recipients were intravenously transplanted into secondary RPK-deficient female mice conditioned with 9 to 9.5 Gy lethal dose irradiation.

When 9 Gy conditioning was applied (assays 1.1), a high survival percentage of receptors was obtained (90%) (Table 13), but no hematological rescue was addressed (Figure 46.A). We observed a modest reduction of splenomegaly in PKD animals transplanted with coRPK expressing cells when comparing with transplanted control animals (Figure 46.B and C) and reticulocytosis was improved at 30 dpt. However, reticulocyte percentage progressively increased with the time reaching levels typical from RPK deficient animals (Figure 46.A). Probably this loss of correction was caused by a loss of donor engraftment, as the number of proviral integrations in blood cells fell down from 4.76 ± 0.28 copies/cell to 1.41 ± 0.59 copies/cell at 30 and 60 dpt respectively (Table 12).

Assay	Groups	Vector Copy Number (VCN/cell)			transduction %	donor chimerism %
		WBC	total BM	indiv. CFU	provirus⁺ CFUs	SRY⁺ PB cells
1	PKD hPGK-EGFP (n=2)	0.83 ± 0.05	0.42 ± 0.03	0.42 ± 0.00	57.73 ± 12.28	n.d
MOI 1-4	PKD hPGK-coRPK (n=3)	4.76 ± 0.28	3.58 ± 0.34	3.07 ± 0.76	91.08 ± 3.67	n.d
2	PKD hPGK-EGFP (n=6)	4.56 ± 0.50	4.19 ± 1.29	n.d	91.13 ± 4.49*	61.82 ± 3.61
MOI 10	PKD hPGK-coRPK (n=14)	1.65 ± 0.08	0.99 ± 0.13	n.d	65 ± 0.75*	63.66 ± 4.45

Table 11: Molecular parameters measured in transplanted mice from two independent gene therapy assays with the hPGK promoter. Number of integrations per cell was calculated by qPCR in WBC, total BM and individual BM CFUs at 140 dpt, by multiplex detection of proviral Ψ sequence and mouse *Titin* endogenous gene. Transduction percentage was calculated as the number of colonies carrying the integrated provirus over the total of colonies analyzed (EGFP n=10 per mouse; coRPK n=25-30 per mouse). Data marked with * symbol represent the estimated transduction percentage. This percentage was obtained by interpolation in the linear regression $y = 8.98 \cdot x + 50.18$ ($R^2 = 0.8196$) built from experiment 1 (X axis: VCN/WBC, Y axis: % provirus$^+$ CFUs). Donor chimerism was calculated at 140 dpt as the % of cells positive for the SRY sequence of the Y chromosome, using a curve of male/female blood cells and normalizing to mouse β-Actin genomic sequences. n.d: not determined.

PRIMARY receptors		SECONDARY receptors							COMMENTS
Assay	Sample	Assay	Cells per mouse	Groups	VCN/cell (WBC)		donor chimerism %		
					30 dpt	60 dpt	30 dpt	60 dpt	
1 MOI 1-4	170 dpt individual frozen BM	1.1	5.10^5-$2.5.10^6$	PKD hPGK-EGFP n=4	0.01 ± 0.01	0.10 ± 0.06	n.d		No hematological correction Lost of reticulocytosis correction at 90 Slight reduction in spleen weight Lost of engraftment
				PKD hPGK-coRPK n=6	4.76 ± 0.28	1.41 ± 0.59			
	170 dpt pooled frozen BM	1.2	5.10^5-3.10^6	PKD hPGK-EGFP n= 4					HEMATOPOIETIC SYMDROME and died
				PKD hPGK-coRPK n=6					
	170 dpt pooled frozen BM	1.3	2.10^6	PKD hPGK-coRPK n=6	3.3				5/6 HEMATOPOIETIC SYMDROME and died correction in one alive mouse
2 MOI 10	140 dpt pooled fresh BM	2.1	4.10^6	PKD hPGK-EGFP n=8	2.47 ± 0.33	1.13 ± 0.20	67.82 ± 14.02	59.86 ± 17.68	4/16 HEMATOPOIETIC SYMDROME and died 8/12 LEUKEMIA development Partial correction (RBC) in surviving mice Low VCN
				PKD hPGK-coRPK n=8	0.81 ± 0.18	0.43 ± 0.14	67.38 ± 7.96	55.35 ± 8.86	
	140 dpt pooled fresh BM	2.2	4.10^6	PKD hPGK-coRPK n=8	1.44 ± 0.08		62.89 ± 5.61		4/8 HEMATOPOIETIC SYMDROME and died Hematological and reticulocytosis correction Reduced splenomegaly Multi-lineage reconstitution

Table 12: Molecular parameters measured in secondary transplanted mice from two independent gene therapy experiments based on lentiviral vectors carrying the hPGK promoter. Number of integrations per cell was calculated in PB at 30 and 60 dpt, as well as donor chimerism percentage. All secondary transplants but assay 1.1, were carried out with 9.5 Gy of irradiation dose. Numbers represent the average per group and SEM.

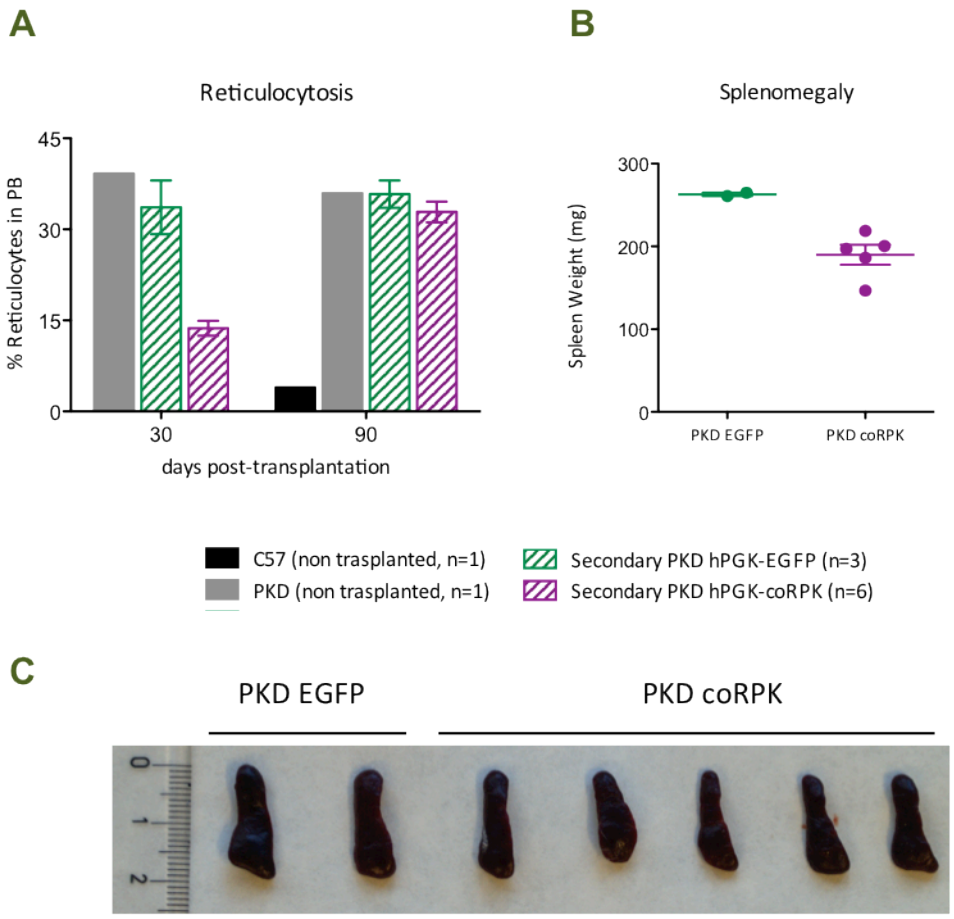

Figure 46: Analysis of PKD secondary receptors from experiment 1.1. **A)** Reticulocyte levels and **B)** spleen weight were measured in secondary recipients transplanted with total BM from primary PKD mice transplanted with pCCL-hPGK-EGFP or pCCL-hPGK-coRPK transduced cells. **C)** Pictures of representative spleen from secondary transplanted mice 90 days after transplantation. Bars represent the average and SEM of positive cells % per group for the different lineages. Data were not statistically analyzed because sample size was < 3 in non-transplanted control groups.

Next sets of secondary transplants were performed increasing the irradiation dose up to 9.5 Gy, which dramatically decreased the survival of receptors. In assay 1.2 all receptors developed hematopoietic syndrome and died soon after conditioning, and only one receptor survived in experiment 1.3. Hence, this mouse showed an increase in the erythrocyte level in PB (6,87.10^9 RBC/mL) with 3.3 proviral integrations per blood cell (Table 12). Experiment 2.1 yielded 75 % survival of receptors after irradiation, which showed normal hematopoietic lineages in PB 2 months after transplantation. However, we observed that some mice exhibited an aberrant expression of hematopoietic markers at 60 dpt (Figure 47) indicating leukaemia development. Nevertheless, the surviving transplanted animals carrying the coRPK transgene maintained the donor chimerism percentage over the time

(estimated at 55.35 ± 8.86 % of SRY^+ cells at 60 dpt) at the same levels as primary receptors, showing a partial long-term hematological correction. Probably this low level of correction was due to the low number of proviral copies detected in blood cells (0.43 ± 0.14 integrations/cell at 60 dpt) (Table 12).

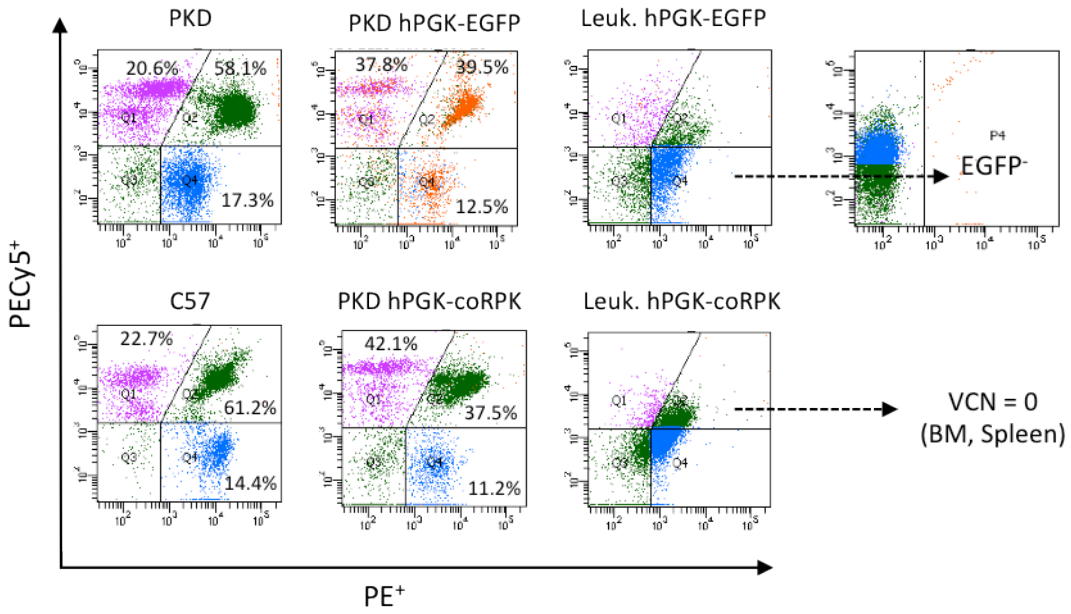

Figure 47: Study of hematopoietic lineages in peripheral blood of secondary transplanted mice from experiment 2.1. Representative dot-plot of the flow cytometry analysis used to detect B and T lymphoid cells using conjugated CD3-PE, B220-PE and B220-PECy5 antibodies, and myeloid cells using Gr-1 and Mac-1 biotinylated anti-mouse antibodies (see Figure 52.A). First and second columns represent normal lineages detected at 60 dpt in all groups of animals. Third column represents the aberrant expression of lineage markers observed in some transplanted mice at 60 dpt. Last column shows that leukemic cells did not express EGFP in the control group, and no proviral copies were detected in secondary receptors from coRPK group.

Additionally, leukemic animals from assay 2.1 were characterized by flow cytometry analysis, detecting B leukemic cells in both groups of transplanted animals but with differential expression of c-kit marker (Appendix 2). Leukaemia onset appeared in 8 out of 12 transplanted mice in this assay, which constitutes only 6.6% of the total transplanted mice (8 out of 121 transplanted mice), or 9% of the surviving mice after conditioning (8 out of 88 mice) (Table 13). Metabolic advantage of WBC expressing the transgenic RPK protein was discarded as the cause of leukaemia because both EGFP and coRPK groups of animals were affected. On the other hand, no proviral integrations were detected in blood cells or spleen from leukemic mice, and leukemic cells did not express EGFP transgene in the control group. These outcomes suggest that lentiviral integration in the host genome was not the cause of leukaemia development in these secondary receptors (Figure 47).

Promoter	Assay	transpanted mice	surviving mice	survival %	Leukemic mice
CMV	1	7	7	100	
	2	9	7	80	
	3	7	7	100	
	Secondary	4	2	50	
	4	8	4	50	
	5	9	9	100	
	Total	**44**	**36**		
hPGK	1	6	6	100	**9%**
	1.1	10	9	90	
	1.2	10	0	0	
	1.3	6	1	16	
	2	21	20	95	
	2.1	16	12*	75	
	2.2	8	4	50	
	Total	**77**	**52**		
TOTAL		**121**	**88**		

Table 13: Number of animals used in each gene therapy experiment. Transplanted animals represent the total number of PKD mice transplanted in each assay, followed by the number of surviving animals after conditioning. The mark (*) indicates that 8 out of 12 surviving mice developed leukaemia, which represents 9% of the total surviving transplanted mice after conditioning (8 out of 88 mice).

In the last set of secondary transplanted mice (assay 2.2), a clear phenotype correction was achieved in animals expressing the coRPK transgene (Figure 48). Surviving receptors (50% of total transplanted mice) showed an improvement of hematological parameters (increased RBC, HGB, HTC and reduction in MCV and MCH) comparable to the phenotypic correction obtained in primary receptors, together with a steady reduction of circulating reticulocytes (Figure 49 and 50). Although the percentage of reticulocytes at 280 dpt in secondary transplanted mice (18.18 ± 3.08 %) was higher than in primary mice at the same time point (11.23 ± 1.54 %) (Figure 36), it constituted just 3% more than the reticulocyte level obtained at the end-point of the primary transplantation assay (15.63 ± 3.20 % at 280 dpt) (Figure 36). Secondary coRPK expressing mice also showed a marked decrease of spleen weight (143.65 ± 12.79 mg), reaching levels closer to wild-type mice (78.15 ± 11.65 mg) and primary corrected transplanted mice (123 ± 9.09 mg), than to PKD anemic mice (297.47 ± 25.54 mg) (Figure 51).

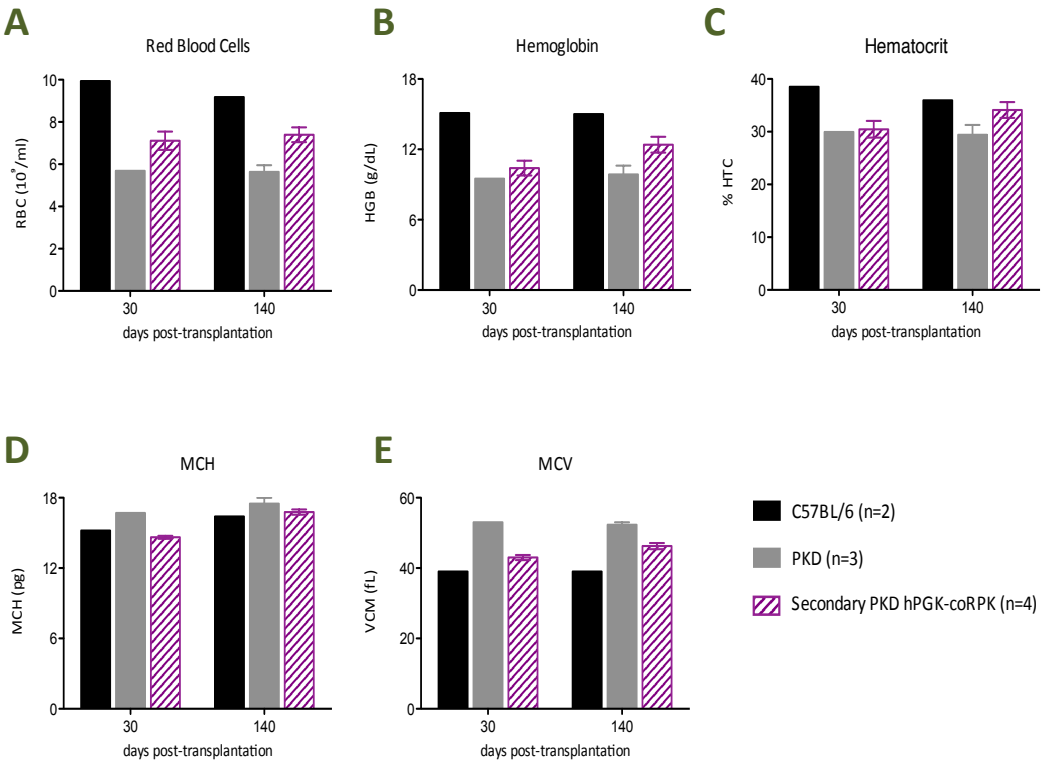

Figure 48: Hematological parameters recorded in secondary receptors from experiment 2.2. **A)** RBC counts, **B)** HGB, **C)** HTC, **D)** MCH and **E)** MCV were measured in PB of secondary transplanted mice at different times post-transplantation, showing a recovery of the hematological phenotype when comparing with RPK deficient controls. Bars represent the average and SEM per group.

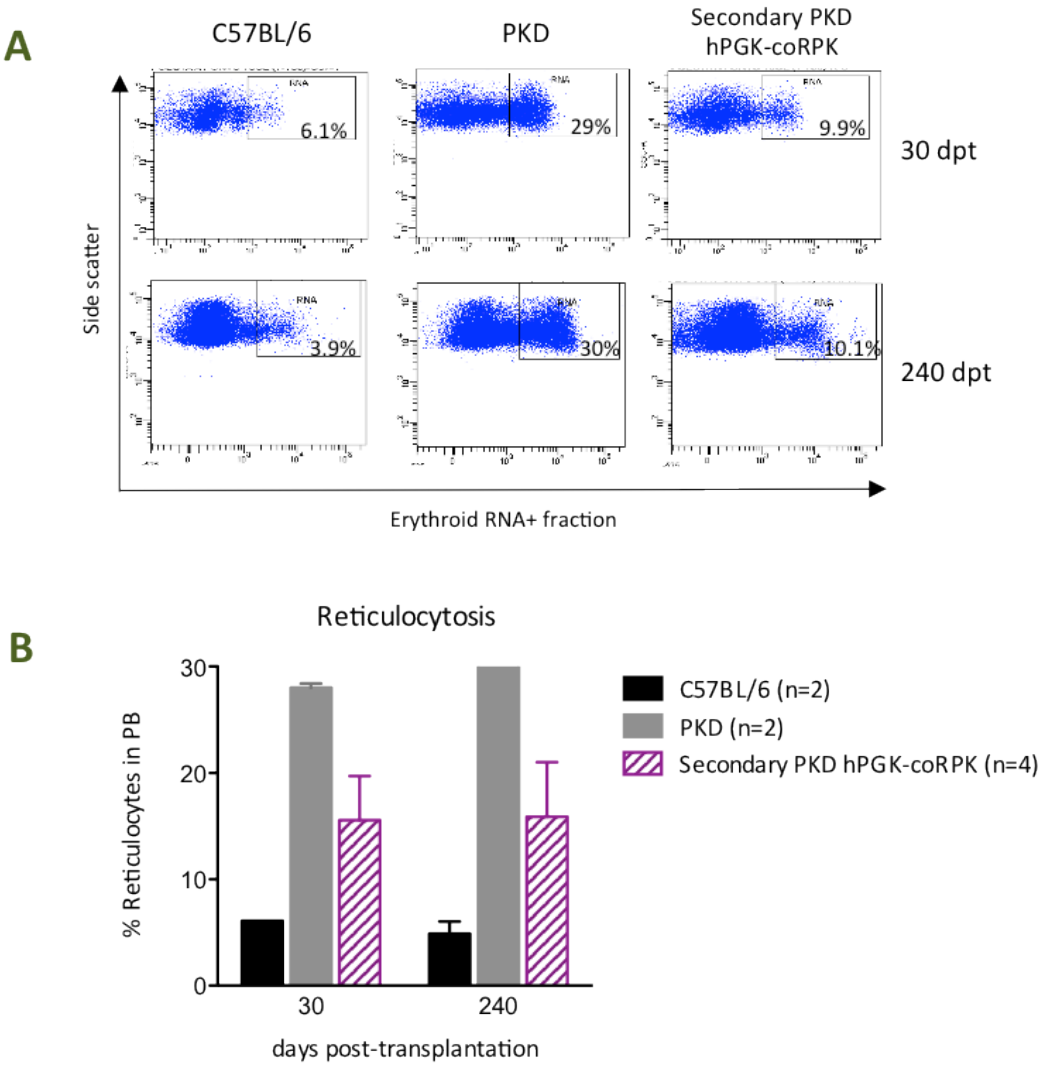

Figure 49: Study of reticulocyte levels in peripheral blood of secondary receptors from experiment 2.2. **A)** Representative dot-plot of the flow cytometry analysis used to identify the reticulocyte population by Acridine Orange staining. **B)** Reticulocyte percentages scored in non-transplanted and transplanted mice at different times post-transplantation. Bars represent the average and SEM per group. Data were not statistically analyzed because sample size was < 3 in non-transplanted control groups.

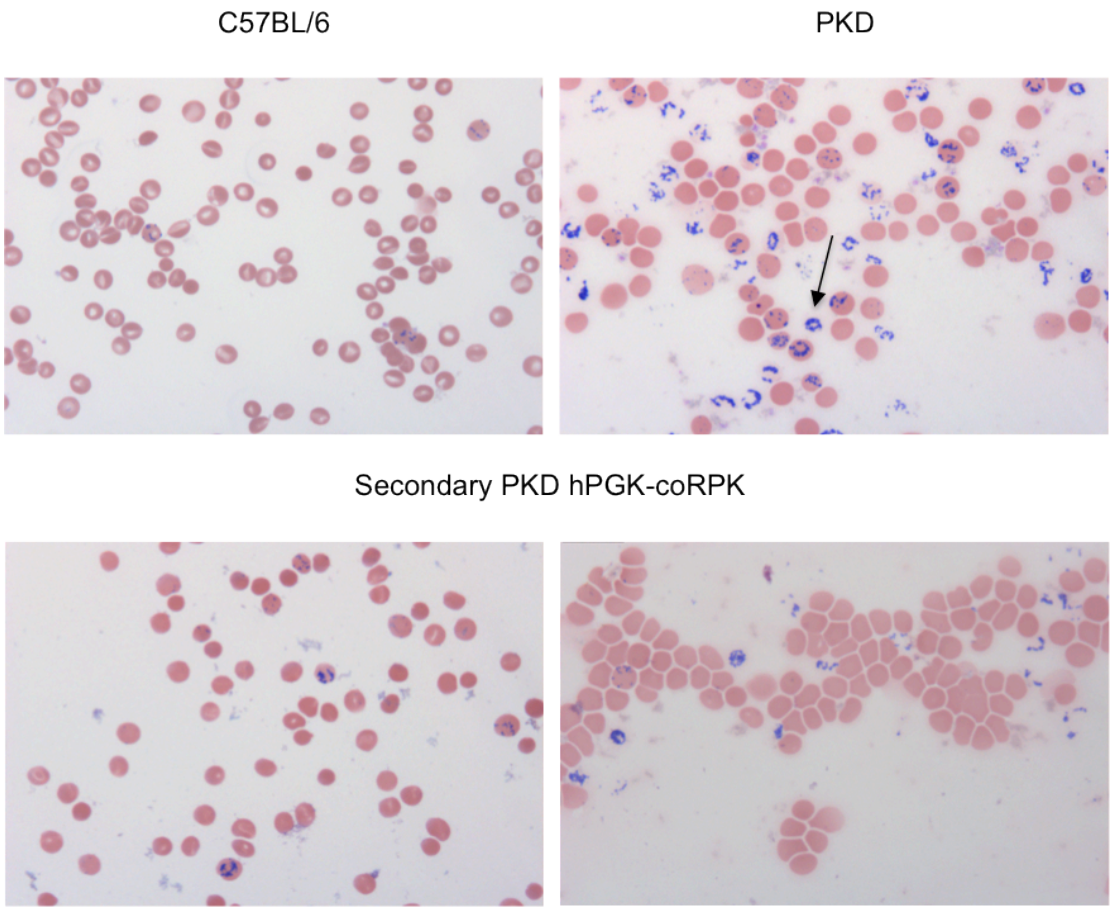

Figure 50: Reticulocyte staining of peripheral blood from secondary receptors (experiment 2.2). Blood smears from non-transplanted and transplanted secondary mice stained with Brilliant Cresyl Blue that binds to RNA detecting the reticulocytes (indicated by an arrow)

Donor chimerism was also determined in peripheral blood from transplanted mice and estimated at 62.89 ± 5.61 %. Both, donor engraftment and the number of proviral integrations per cell (1.44 ± 0.08 provirus copies/cell in PB) were equal to primary receptors (Table 12). Transduced hematopoietic progenitors from primary recipients were able to engraft in secondary receptors, showing a multi-lineage reconstitution in PB up to 5 months post-transplantation (Figure 52.C), with no signs of leukaemia development. Altogether these results indicate that transgenic expression of the human RPK protein completely reversed the hematological symptoms of the PKD in primary receptors, maintaining the stable correction of the phenotype in secondary transplanted mice.

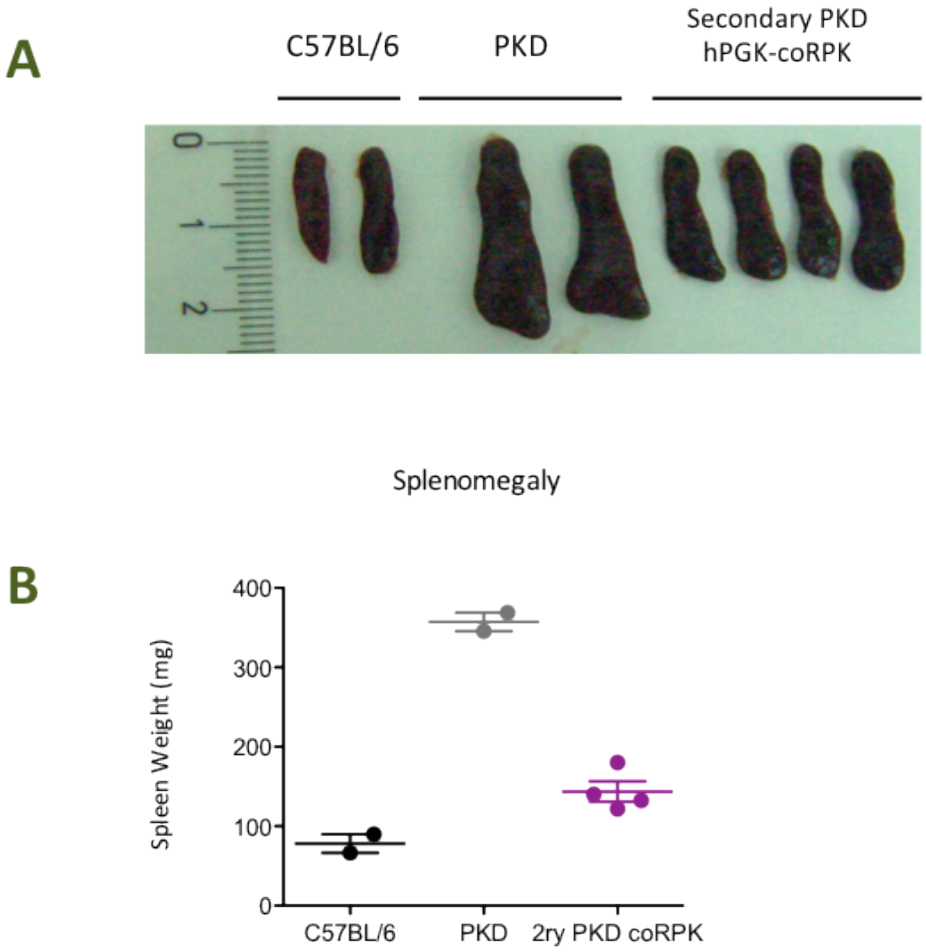

Figure 51: Spleen weight and size of secondary receptors from experiment 2.2. **A)** Pictures of representative spleen from secondary transplanted mice 240 days after transplantation. **B)** Spleen weight of non-transplanted and secondary transplanted mice. Bars represent the average and SEM per group.

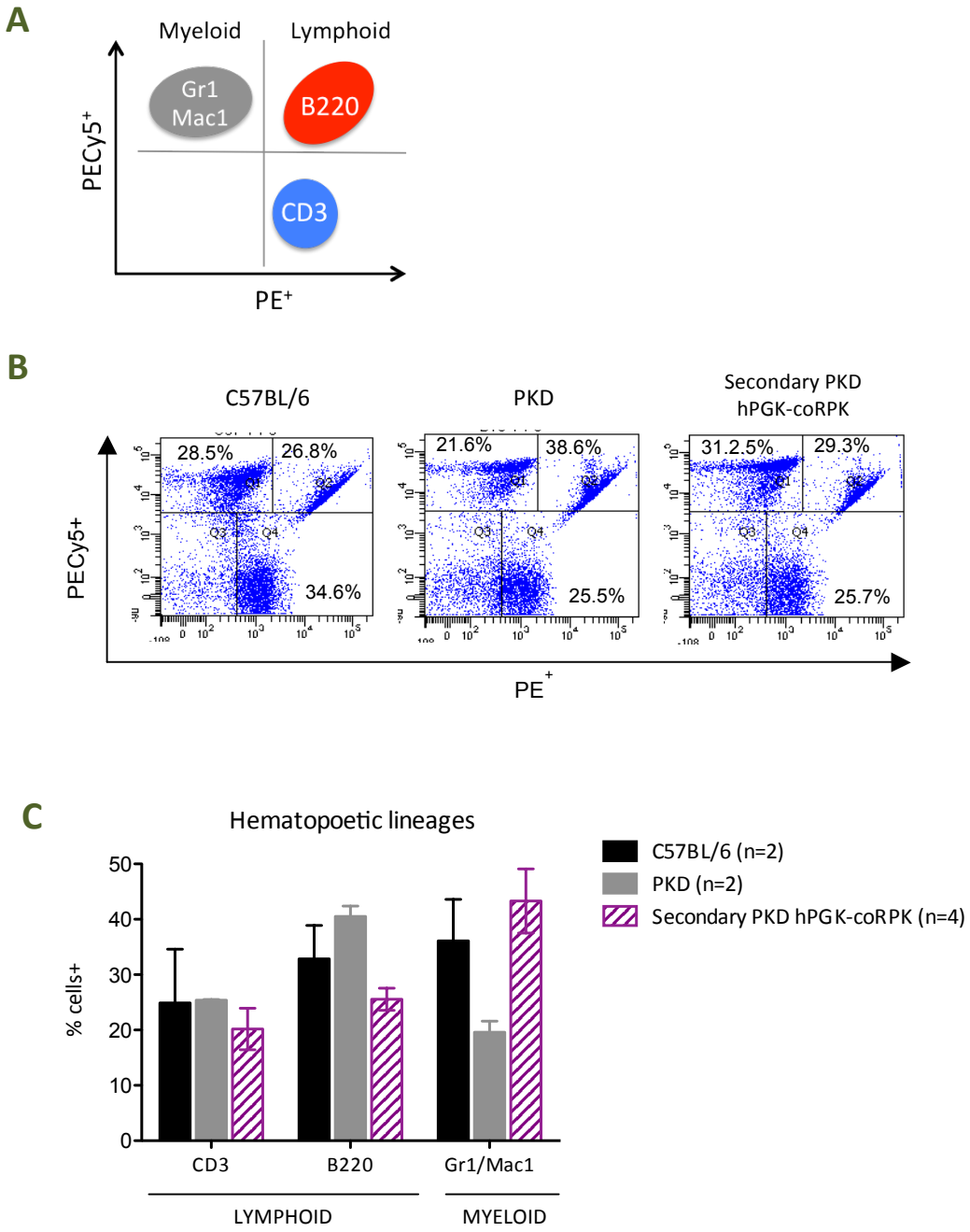

Figure 52: Study of hematopoietic lineages in peripheral blood of secondary transplanted mice from experiment 2.2. **A)** Diagram of the different hematopoietic lineages identified by labeling with CD3-PE, B220-PE, B220-PECy5, Gr1-Biotin and Mac1-Biotin antibodies. **B)** Representative dot-plot of the flow cytometry analysis used. **C)** Percentages of each lineage in PB at 140 dpt. Bars represent the average and SEM of positive cells % per group.

2.2.7 Metabolomic studies in primary PKD transplanted mice

To study whether the ectopic expression of RPK was able to correct the metabolic perturbations observed in RPK deficient animals, we performed an extensive metabolomic analysis of all transplanted and control mice. Untargeted profiling technology was applied to RBC and WBC samples, revealing significant changes in the RBC expression profile of metabolites involved in the glycolysis pathway, by comparison of all groups to PKD mice. We found three broad clusters with distinct trends of expression pattern (Figure 53) among the different groups.

RBC expressing coRPK transgene showed an increase in metabolite expression from cluster 1 similar to normal C57BL/6 controls as well as a decrease in cluster 2 metabolites similar to normal mice and distinct from PKD or EGFP controls. In the third cluster, however, metabolite profile showed an increment pattern similar to mice carrying the EGFP control transgene (assay 2) (Figure 53), which may be related with the lentiviral integration itself. On the other hand, differences found in these three metabolic profile clusters were more obvious in the first experiment than in the second, probably due to the transduction efficacy obtained or the cohort size used in each one. Additionally, two samples from AB6F1 control mice were included in the study, showing a metabolite pattern of expression very similar to C57BL/6 mice, and therefore very similar to genetically corrected PKD animals in the first two clusters. We also analyzed by targeted LC/MS other selected glycolytic metabolites that did not score as significant, but we obtained low confidence in these studies. Even so, metabolic profiling showed some expression trends in corrected PKD animals and C57BL/6 wild-type mice, such as an increase in ATP and ADP, and a decrease in alanine (data not shown).

In a second metabolomic study with higher sample sizes and using A/J wild-type mice as controls to avoid differences caused by genetic background, untargeted profiling showed similar trend as before (Figure 54). PKD mice expressing the transgenic RPK protein compared to PKD deficient animals yielded a similar profile as the comparison between PKD animals and A/J mice, but not when comparing with mice carrying the EGFP control transgene.

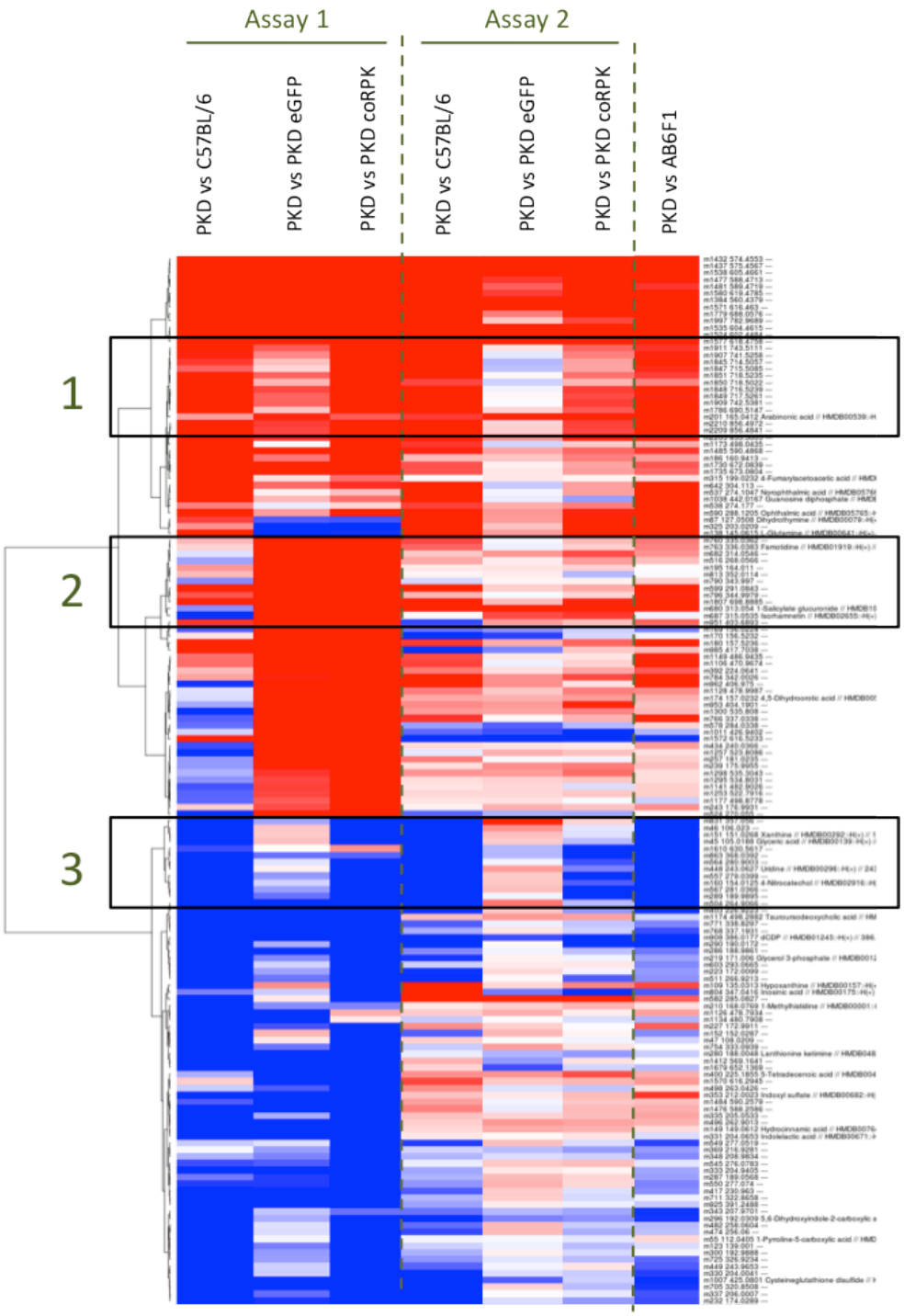

Figure 53: Analysis of significant metabolic profile changes by comparing to PKD deficient animals in two independent experiments. Data obtained from complete RBC heat map (E921), where higher and lower expressions are represented in red and blue respectively. Metabolites listed have at least one comparison that is significant using the following criteria: absolute fold change>1.5; minimal signal>2000; Adjusted p-value< 0.01. Black boxes highlight cluster of metabolite changes with distinct profile among the groups. Assay 1: C57BL/6 n=1, PKD n=1, PKD hPGK-EGFP n=2, PKD hPGK-coRPK n=3. Assay 2: C57BL/6 n=2, PKD n=2, PKD hPGK-EGFP n=6, PKD hPGK-coRPK n=10, AB6F1 n=2.

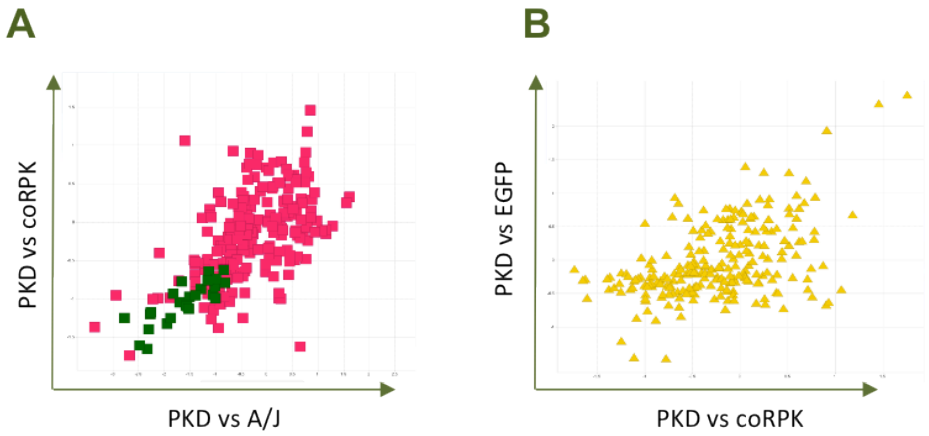

Figure 54: Untargeted profiling analysis of selected number of metabolites involved in the glycolytic pathway from control and transplanted mice RBC. **A)** Metabolite expression profile comparison between A/J wild-type mice or **B)** PKD animals carrying the EGFP reporter transgene, and genetically corrected PKD animals in relation to metabolites that have at least one significant comparison using the following criteria: absolute fold change>1.5; minimal signal>2000; Adjusted p-value< 0.01. A/J n=3, PKD n=7, PKD hPGK-EGFP n=11, PKD hPGK-coRPK n=18.

Targeted LC/MS profiling performed on RBC samples from all groups revealed a few metabolites that changed in a predicted manner, such as phosphoenolpyruvate (PEP) and 3-phosphoglycerate (3-PG), whose expression approached to C57BL/6 or A/J wild-type mice expression profile (Figure 55.A and B). However, many other metabolites did not show a discernable pattern, such as pyruvate (Figure 55.C) or fructose 1,6-diphosphate, and signal from 2,3-Bisphosphoglycerate (2,3-BPG) was saturated (data not shown). Additionally, we did observe that genetically corrected mice showed an increased trend of D-Lactate and two metabolites involved in pentose phosphate pathway (D-Ribose-5-phosphate and D-Ribulose-5-phosphate) compared to all analyzed groups (data not shown).

Because of the possibility of transgenic expression of RPK using hPGK ubiquitous disturbing physiological ATP production in lineages other than erythroid, with the consequent aberrant metabolic alterations, metabolic profile of WBC was analyzed in parallel to RBC for all mice. As shown in Figure 56, metabolite expression pattern of RBC was different depending on the groups and was markedly different to WBC profile. On the contrary, WBC sub-groups clustered together with very little difference among groups, which indicates no metabolic balance alterations in leukocytes with ectopic expression of the coRPK transgene.

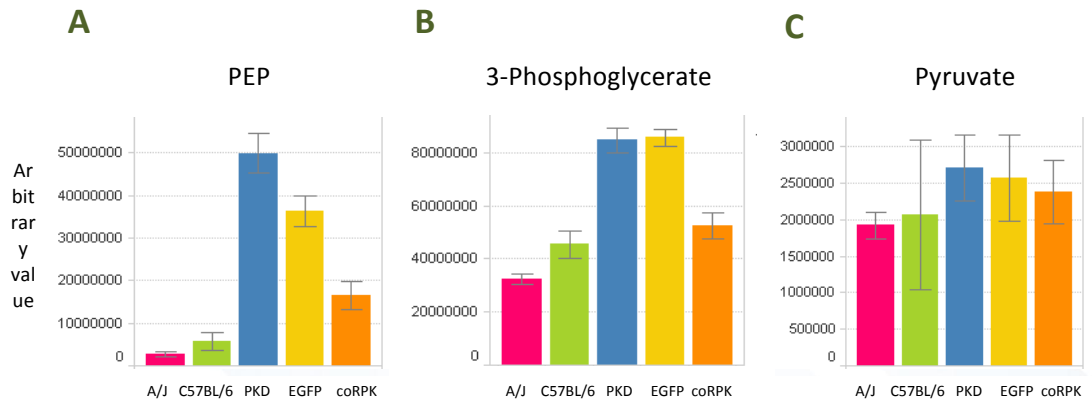

Figure 55: Analysis of selected metabolites in RBC by targeted LC/MS profiling. Data represent the average and SEM of metabolite expression per group from both experiments together. **A)** Phosphoelopyruvate (PEP) and **B)** 3-phosphoglycerate (3-PG) showed a decreased expression trend, but **C)** pyruvate did not significantly change comparing to anemic mice. A/J n=3, PKD n=7, PKD hPGK-EGFP n=11, PKD hPGK-coRPK n=18.

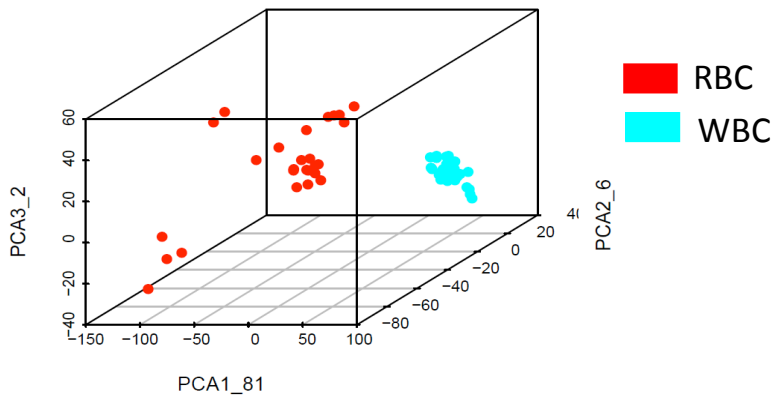

Figure 56: Untargeted profiling analysis of selected number of metabolites involved in the glycolytic pathway from WBC and RBC samples of control and transplanted mice. Metabolite expression profile comparison among C57BL/6 (n=3), AB6F1 (n=2) and PKD (n=3) control groups and PKD mice transplanted with pCCL-hPGK-EGFP (n=8) or pCCL-hPGK-coRPK (n=13) transduced cells.

Overall these results demonstrated that ectopic expression of RPK largely recapitulated the metabolic profile of wild-type RBC independently of the wild-type mouse strain used for comparing (C57BL/6 or A/J), correcting the energetic defect in RBC without changing the metabolic balance in WBC.

3. Erythroid Tissue-Specific Lentiviral Vectors for the PKD

We have shown above that human PGK promoter within a SIN lentiviral backbone was able to drive a stable and long-term coRPK expression, achieving an efficient correction of the PKD anemic phenotype in the AcB55 mouse model. Although the metabolomic studies showed that the gene therapy treatment applied did not modify the metabolic balance in leukocytes, the restriction of the transgene expression to the affected cell type could emerge as a safer approach for clinical gene therapy applications. Through the use of tissue-specific promoters/enhancers, the potential effects of the integrative vectors in the host genome will be decreased, reducing the risk of insertional mutagenesis. This is particularly important when the gene therapy treatment involves reconstitution of HSCs.

To take a step forward in the gene therapy for PKD, we developed new SIN lentiviral vectors carrying the regulatory sequences of the human *PKLR* gene driving the expression of the EGFP or the coRPK transgenes (Figure 15). However, in preliminary experiments PKLR promoter provided low levels of transgene expression, so we included the *cis*-acting regulatory elements of the β-globin LCR to increase the expression of the transgene (Figure 16). These regulatory sequences of the β-globin gene have been seen to increase the erythroid-specific expression in transgenic mice (May et al., 2000) (May et al., 2002) and improved the safety profile in *in vitro* immortalization assays (Arumugam et al., 2009). We also developed a control erythroid vector harbouring the β-globin promoter and the βLCR sequences (Figure 16) driving the expression of the coRPK transgene. This combination of promoter and regulatory elements was previously validated in a lentiviral backbone in a gene therapy pre-clinical model for β-thalassaemia in mice (Miccio et al., 2008) and it is being currently used in a clinical trial for β-thalassemia with successful results in one adult patient (Cavazzana-Calvo et al., 2010).

First, we checked the erythroid specificity of the human PKLR promoter by comparison with the ubiquitous human PGK promoter, both regulating the expression of EGFP reporter gene. Myeloid (HL60) and erythroid (K562 and HEL) human cell lines were transduced with each vector in parallel, using the same MOI and number of cells, obtaining high level of transgene expression in all cell lines when using hPGK promoter (Figure 57 and

58). However, when expression was driven by the PKLR promoter, EGFP expression was only detected in K562 and HEL erythroid cell lines at low levels (13-15%), but not in HL60 non-erythroid cell line, confirming the specificity of the human *PKLR* gene regulatory sequences.

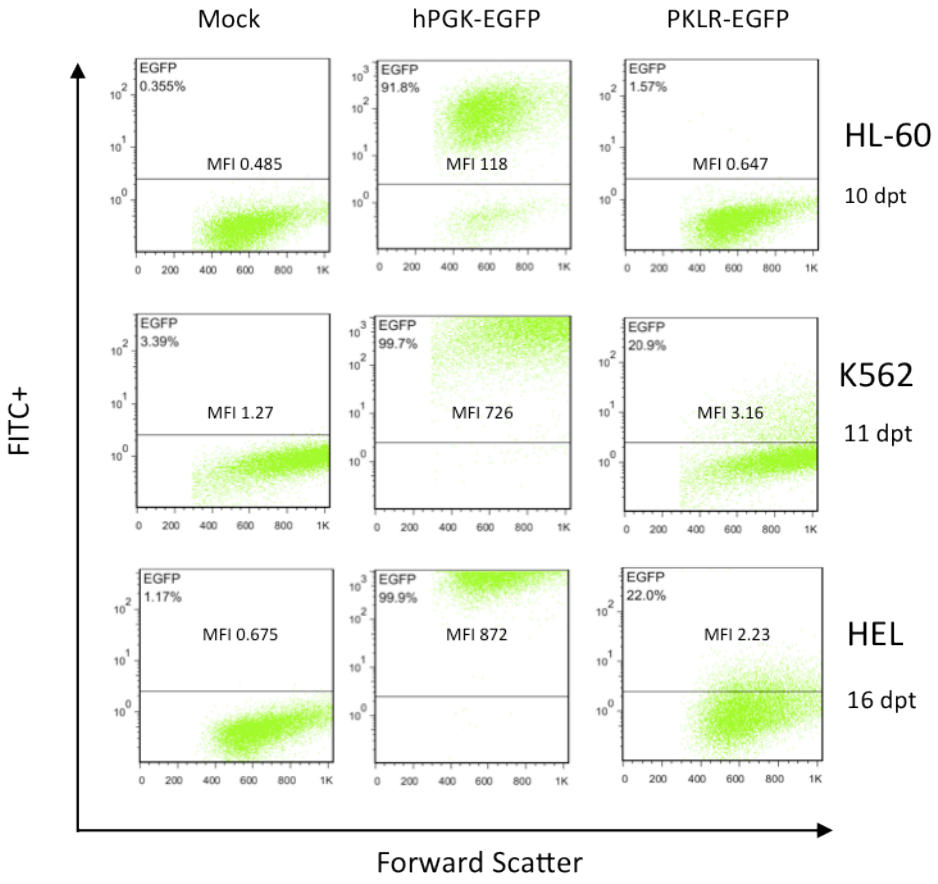

Figure 57: Flow cytometry analysis of different human cell lines transduced with lentiviral vectors carrying the hPGK or PKLR promoter. Myeloid HL60 cell line and K562 and HEL human erythroid cell lines were infected at MOI 4 with lentiviral vectors harbouring the EGFP transgene and either hPGK or PKLR promoter. Same non-infected cell lines were maintained in parallel as controls. 10 to 16 days after transduction, cells were harvested and EGFP expression was analyzed.

In order to enhance the transgene expression driven by the erythroid-specific PKLR promoter, we included the βLCR regulatory sequences to the expression cassette. HS elements from βLCR are nucleosome-free regions of open chromatin that are highly accessible to *trans*-acting factors, and therefore confer silencing modulation activity. HS4, HS3 and HS2 sites were added to PKLR-coRPK lentiviral construction (Figure 16), and a control lentiviral construction carrying the β-globin promoter and coRPK transgene as well as the HS3 and HS2 sites (Figure 16) were also used, because this combination of enhancers

and promoter had been previously tested. The same cell lines as before were transduced with the different lentiviral vectors and human RPK protein was detected by western blot. As figure 58 shows, K562 was the only cell line that expresses the endogenous RPK protein when no differentiation protocols are applied, because RPK was detected in non-transduced K562 cells. PGK promoter drove the ubiquitous expression of the transgene, detecting human RPK in all cell lines infected with this vector (Figure 58) at similar levels of expression (Figure 59). On the contrary, both PKLR and β-globin promoter showed high erythroid specificity, because we only detected human RPK protein in K562 and HEL erythroid cells

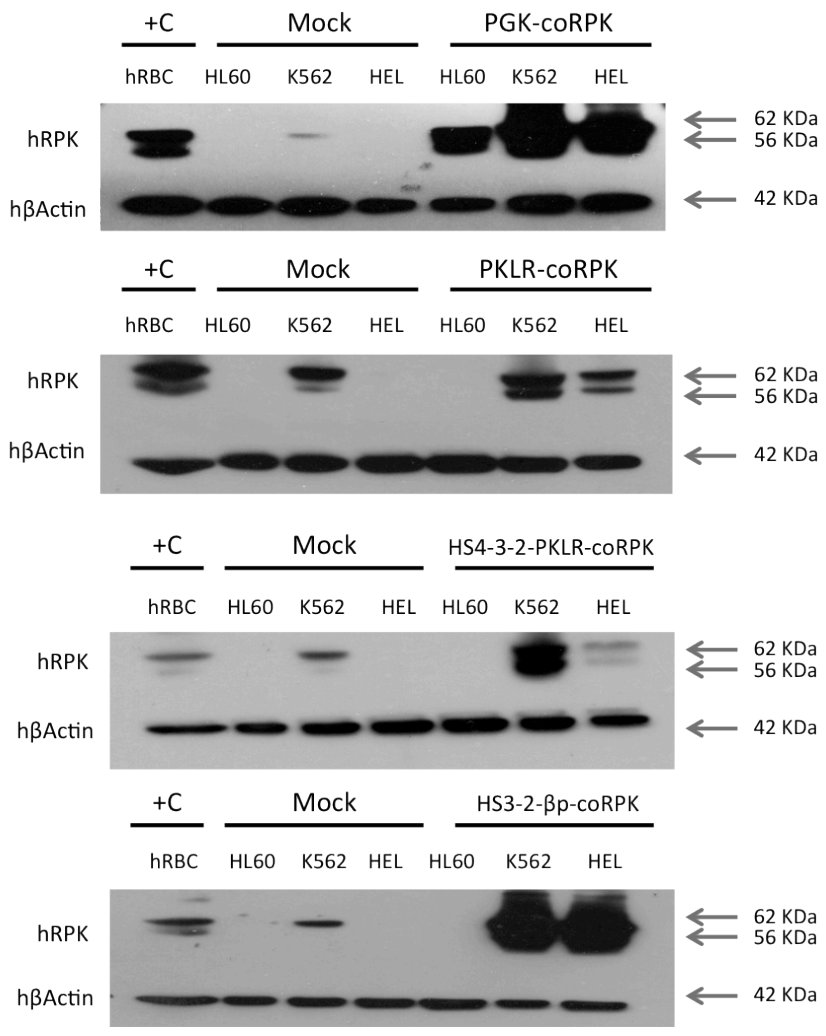

Figure 58: Western blot analysis of different cell lines transduced with the developed lentiviral vectors carrying the coRPK transgene. Total protein extracts were obtained from HL60, K562 and HEL cells 12-20 days after infection with PGK, PKLR, HS4-HS3-HS2-PKLR or HS3-HS2-βp- promoter carrying lentiviral vectors at MOI 4. Samples were analyzed by Western blot for human RPK protein and mouse β-Actin as loading control. Human peripheral blood lysates were used as positive controls and non-transduced cells were carried in parallel as endogenous RPK expression control.

infected with these vectors, even at long exposure times (Figure 58). However, the quantification of the band intensities revealed that β-globin promoter conferred stronger expression of the coRPK transgene than the PKLR promoter even when we introduced the HS enhancers of the βLCR (2.3-fold and 1.8-fold more expression in K562 and HEL cells respectively) (Figure 59). Although the three lentiviral vectors developed were highly specific for the erythroid lineage, HS3-HS2-βp combination was the most efficient one driving the coRPK expression, reaching even higher expression in erythroid cells than the hPGK promoter.

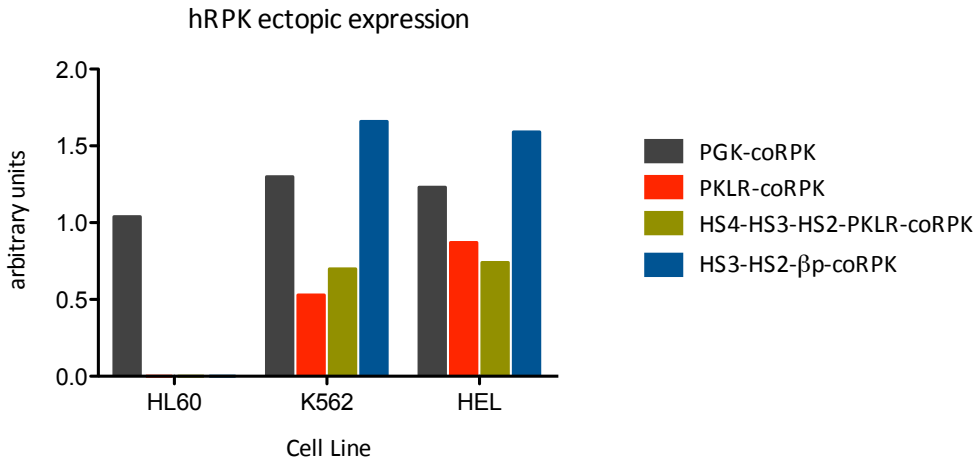

Figure 59: Quantification of the bands obtained by Western blot corresponding to human RPK protein from the different transduced cell lines. Human RPK expression levels were quantified by densitometry and normalized to murine β-Actin expression. Bars represent the ectopic expression of the RPK driven by the different enhancer/promoters removing RPK endogenous expression in K562 cells.

VII. DISCUSSION

1. AcB55 mouse strain is a suitable model of human PKD disease

Red cell pyruvate kinase deficiency is the most frequent hereditary enzymatic defect of the glycolysis pathway in erythrocytes, and causes CNSHA. More than 400 cases of PKD have so far been reported. From a clinical point of view, PKD is a complex disease because of its high molecular heterogeneity and variable symptomatology. These characteristics hamper accurate diagnosis of PKD and the choice of an appropriate treatment. In this sense, the availability of mouse models that reproduce the human PKD symptoms allows not only the study of the pathological mechanisms involved, but also to explore new therapeutic options.

Unlike other erythroid diseases, such as thalassaemia major or SCD, PKD does not have a transgenic animal model, but several *Pklr* wild mutant mice and dogs have been identified. The recombinant congenic AcB55 mouse strain was the latest of being found (Min-Oo et al., 2003), and has been the experimental model used in this work. We first checked that AcB55 mice reproduced the PKD hematological phenotype because these mice had not been previously characterized in detail, identifying the presence of splenomegaly, reduced numbers of RBCs and HGB, constitutive reticulocytosis, shortened RBC half-life and extramedullar erythropoiesis in spleen and liver.

Our results showed, just like Min-Oo et al. (Min-Oo et al., 2004), that AcB55 mice reproduce the anemic status of PKD patients. We found that 269 T>A loss-of-function mutation in the *Pklr* gene led to a reduction in the main hematological parameters, including RBC counts, HTC percentage and HGB levels, as well as an increase of MCV and MCH (Figure 18) causing severe hemolytic anemia. The 40% reduction of circulating RBC found in PKD animals (Table 7), together with the drop of hematocrit percentage, indicated a quick loss of deficient RBCs. In addition, RBC volume (MCV) in PKD mice was 25% higher than in controls (Table 7), probably related to the cell volume expansion during the hemolytic process. Reticulocytosis and splenomegaly, typical symptoms of PKD, were also markedly present in PKD mice, showing increases of over 300% with respect to the control mice (Figures 19 and 20). These findings indicated that non-functional RBCs were quickly removed from circulation, triggering a pathophysiological adaptation to release reticulocytes from the BM

to peripheral blood, as has been described for most PKD patients (Diez et al., 2005); (Lakomek et al., 1989) and for CBA PK-1^{slc}/Pk-1^{slc} deficient mice (Kanno et al., 1995). Due to immature erythroid cells still keep some mitochondrial activity (Keitt, 1966), reticulocytes are likely to be the cells responsible for AcB55 mice survival even in an acute hemolytic situation, covering RBC metabolic demands.

This rapid replenishment of damaged RBC also agrees with the increased spleen size observed in RPK deficient animals (Figure 20). The same mechanism that spleen macrophages use to recognize senescent RBCs may be acting in PKD disease. The morphological changes determining RBC destruction in the spleen are not entirely clear, but reduction of erythrocyte deformability and surface protein alterations, such as phosphatidylserine exposure, have been involved in the short survival of RBC during hemolytic anemia (de Jong et al., 2001) (Boas et al., 1998); (Connor et al., 1994). In addition, the increased oxidative stress found in RBCs from some PKD patients (Zerez and Tanaka, 1987) together with the expected impairment of K^+ pumps and the increased intracellular calcium levels in ATP-limiting conditions (Miwa and Fujii, 1985), support the hypothesis that deficient RPKs could cause alterations in RBC membrane and cell density, leading to early RBC destruction in the spleen (Jacobasch and Rapoport, 1996) (Figure 60). Interestingly, both increased oxidative stress and membrane alterations have been proposed as possible mechanisms mediating malaria resistance in this PKD mouse model (Min-Oo et al., 2004).

Comparative studies performed by Min-Oo et al between the two available PKD mouse models, AcB55 and CBA PK-1^{slc}/Pk-1^{slc}, pointed to a more deleterious mutation in the CBA PK-1^{slc}/Pk-1^{slc} strain due to lower RPK enzymatic activity and more severe hemolytic anemia in CBA PK-1^{slc}/Pk-1^{slc} mice (Min-Oo et al., 2007). However, our phenotype characterization revealed similar hematological parameters in both PKD strains (RBC, HTC, HGB) and a similar increase of spleen weight over the controls (3 and 4.6 times in CBA PK-1^{slc}/Pk-1^{slc} and AcB55 mice respectively). Reticulocytosis was only 10% higher in CBA PK-1^{slc}/Pk-1^{slc} (Morimoto et al., 1995) than in AcB55 mice (41.6 ± 1.5% and 30.85 ± 0.84%, respectively), but the different method used can account for these differences. In addition, our study included new features not studied in these models before. Thus, similar to PKD patients, different *Pklr* mutations carried by CBA PK-1^{slc}/Pk-1^{slc} and AcB55 mice likely account for these differences in hemolytic phenotype, underlining the molecular and symptomatic heterogeneity of the disease (Berghout et al., 2012).

The study of RBC survival also demonstrated that the metabolic impairment caused by the *Pklr* mutation in AcB55 mice led to a shortened RBC lifespan, displaying a 1.5 shorter half-life than C57BL/6 RBCs (Figure 21). In general, our red cell turnover results agree with previously reported data using the same control mice (Marinkovic et al., 2007) and a similar PKD mouse (AcB61, (Min-Oo et al., 2007) with the same labelling method. However, studies from Min-Oo showed a more dramatic shortening of RBC half-life on CBA PK-1^{slc}/Pk-1^{slc} mice (6 times shorter than $CBA^{+/+}$ controls) (Morimoto et al., 1995) than in AcB61 mice (Min-Oo et al., 2007). Nevertheless, in our C57BL/6 control RBC half-life did not reach the $CBA^{+/+}$ control values reported, suggesting that the genetic background may also determine RBC turnover. This idea is also supported by different RPK activity found in different inbred mouse strains (Bulfield et al., 1978). In addition, different labelling methods used and different degrees of reticulocytosis can account for these discrepancies because none of these methods distinguish between labelled RBCs and labelled reticulocytes.

The shortened RBC half-life observed in AcB55 deficient animals can be explain by both energetic deficit and membrane alterations. Due to erythrocytes being totally dependent on glycolysis for ATP production, disturbances in RBC energetic metabolism caused by mutations in the *Pklr* gene dramatically contribute to shortened RBC half-life. The accumulation of glycolytic intermediates (control mice in Figure 55) located upstream of the RPK-catalysed reaction, observed in both PKD patients (Zanella et al., 2005) and CBA PK-1^{slc}/Pk-1^{slc} deficient mice (Morimoto et al., 1995), ultimately causes RBC premature death due to metabolic block. In addition, although increased 2,3-DPG levels observed in PKD patients (Lakomek et al., 1994) contribute to their survival, this increment can inhibit G6PDH (Lakomek et al., 1994) and hexokinase (Zanella and Bianchi, 2000) enzymes, impairing the glycolytic flux and reducing NADPH in deficient RBCs. Reduced levels of NADPH are especially harmful, as both the glutathione cycle and the pentose-phosphate pathway can be hindered (Zerez and Tanaka, 1987) (Figure 60), likely altering the antioxidant mechanisms of PKD RBCs to detoxify oxygen reactive species (ROS) (Figure 1).

Circulating erythrocytes are exposed to some of the highest levels of oxidative stress in the body, as they carry oxygen bound to hemoglobin (Winterbourn, 1990) and they are exposed to free iron, increasing the content of reactive oxygen species (ROS) in mature RBCs. This suggests that PKD RBCs may be more sensitive to oxidative damage, contributing

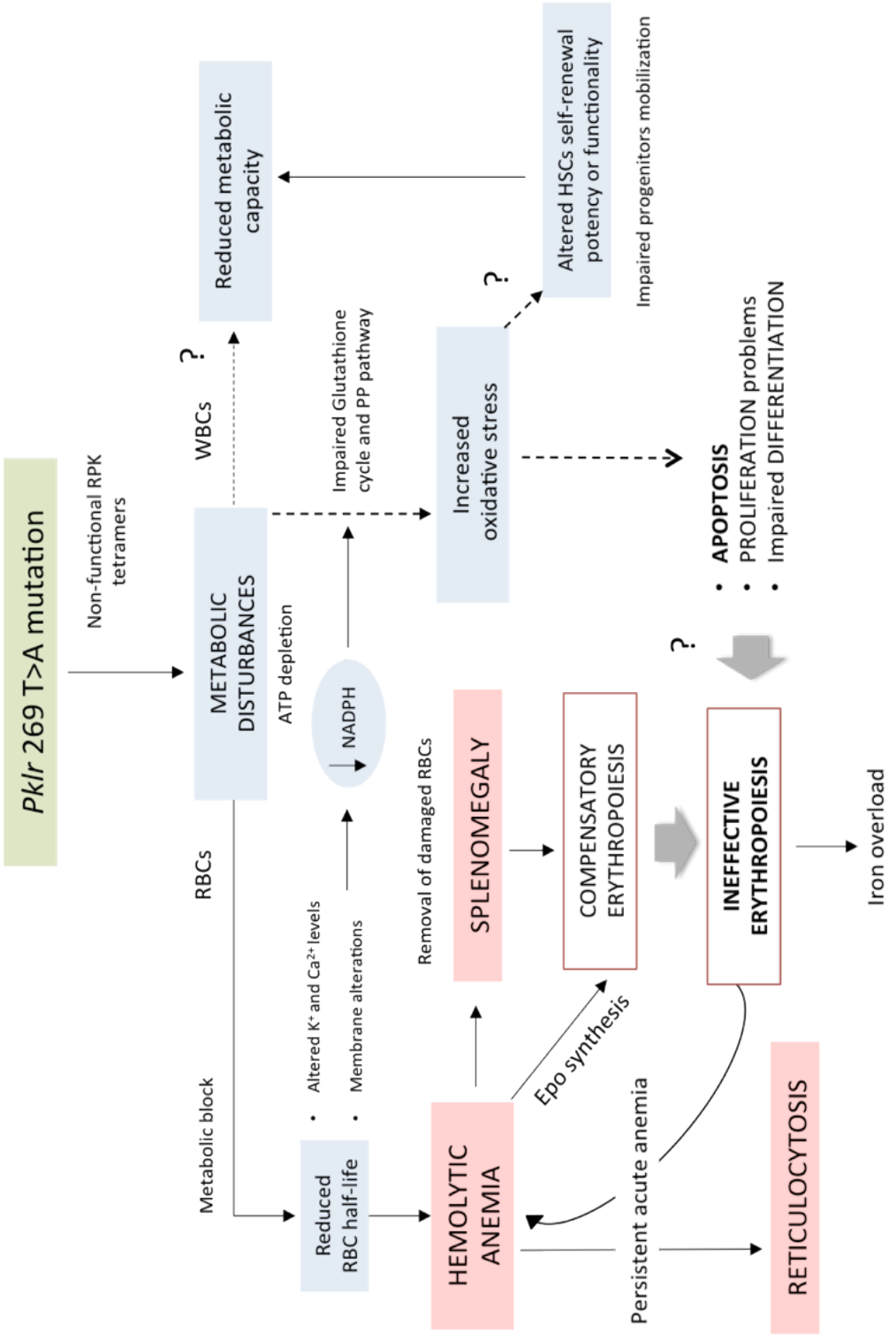

Figure 60: Schematic representation of PKD hematological pathology. Figure shows the hematological parameters studied in the PKD mouse model, as well as possible causes and/or consequences of metabolic disturbances caused by non-functional RPKs. Dashed arrows indicate open questions emerged from this study.

to their shortened survival in the bloodstream. *In vitro* transfection studies from Aisaki et al (Aisaki et al., 2007) and the more severe hemolytic anemia observed in PKD patients suffering infections, oxidative or metabolic stress (Beutler et al., 1978); (Beutler et al., 1986) reinforce this hypothesis. We have started the study of oxidative stress in AcB55 mice to address if these PKD mice also reproduce the altered response to oxidative stress observed in PKD patients. Preliminary data seem to support this hypothesis.

At the same time, impairment of RPK-catalysed reaction likely affects other processes in the cell that require energy, such as membrane protein phosphorilation (Jacobasch and Rapoport, 1996). As mentioned above, the imbalance of K^+ and Ca^{2+} intracellular levels in PKD RBC (Miwa and Fujii, 1985) as a consequence of the ATP-limiting conditions, have been associated with increased apoptosis (Tsujimoto, 1997). On the other hand, a reduction in membrane protein phosphorylation caused by ATP depletion (Jacobasch and Rapoport, 1996) would explain the low initial percentage of labelled-RBCs obtained in PKD animals compared to C57BL/6 controls (Figures 21 and controls in Figures 37 and 38). This hypothesis is also supported by the observation that high levels of intracellular Ca^{2+} in RBCs from PKD patients were involved in Band 3 degradation, a key protein in the erythrocyte membrane structure, through the activation of μ-calpain protease (Passow, 1986).

Through histological analysis and flow cytometry characterization we also demonstrated that AcB55 mice reproduce the CNSHA phenotype observed in PKD patients. Anemic mice showed remarkable splenomegaly (Figure 20) and an expansion of the erythroid compartment in both spleen and liver, indicating the presence of an intense extramedullar erythropoiesis (Figures 23 and 24). In addition, we observed that PKD mice displayed acute iron overload in the liver as a consequence of the hemolytic process, reproducing one of the major problems in severe PKD patients (Marshall et al., 2003) or in repeatedly transfused patients (Wang et al., 2001). Nonetheless, the differences between human and mouse erythropoiesis require careful extrapolation of the data. In humans, normal erythropoiesis takes place in the BM, and only under special circumstances, such as acute hemolysis or stress erythropoiesis, the spleen can serve as an erythropoietic organ. For instance, in PKD patients, splenomegaly is attributed to intense red cell sequestration rather a place for ectopic erythropoiesis and, in general, the clinical status of the patients improves after splenectomy. However, this clinical procedure is not recommended for every

patient, and the clinical benefit mainly depends on individual differences of splenic function (Zanella et al., 2005). On the other hand, splenectomy in CBA mice significantly decreases their lifespan (Aizawa et al., 2005), underlining the key role of spleen in PKD mouse pathology. This may be another reason why the PKD phenotype in humans is more severe than in mice.

The altered erythroid differentiation pattern found in bone marrow and spleen of RPK deficient mice (Figures 22 and controls in Figure 43) also demonstrated that constitutive hemolysis induced an increased erythropoietic response to the acute anemia in PKD mice, being more evident in the spleen. PKD anemic mice showed a predominance of immature erythroid precursors in BM and spleen (proerythroblasts and basophilic erythroblasts) (Figure 22), as well as increased Epo levels in plasma (controls in Figures 31 and 45), because the hematopoietic system attempts to compensate the reduced blood capacity for oxygen transport. However, this stress erythropoiesis was unable to increase the number of mature RBCs, as mature RBCs (IV subpopulation) were clearly reduced in PKD animals, even under high levels of Epo and circulating reticulocytes.

These results demonstrated that *Pklr* mutations led to ineffective erythropoiesis, mainly blocking the step from basophilic erythroblasts (II) to more mature erythroid cells (polychromatophilic and orthochromatophilic erythroblasts) (III and IV) (Figure 22), and therefore failing to produce mature RBCs. These results coincide with the increase of RPK isoform expression over the PKM2 isoform throughout erythroid maturation (Nijhof et al., 1984), being higher in basophilic and orthochromatophilic erythroblasts, and suggest a key role of RPK in late erythroid progenitors maturation. This conclusion is also supported by the increased apoptosis observed by Aizawa et al in the spleen erythroid compartment of both CBA PK-1^{slc}/Pk-1^{slc} mutant mice (Aizawa et al., 2005) and in PKD patients (Aizawa et al., 2003). As mentioned before, impaired antioxidant mechanisms of RPK deficient erythroid cells may become a serious problem for erythroid cells to avoid apoptosis. However, additional causes may contribute to ineffective erythropoiesis in this PKD model. For instance, RPK defects may lead to an impaired proliferation, caused by an exhaustion of erythroid progenitors as a consequence of the stress erythropoiesis. Otherwise, erythroid precursors may continue to proliferate but fail to differentiate, limiting the production of mature RBC.

The hemolytic process in PKD mice renders high serum/plasma levels of Epo (controls in Figures 31 and 45) because the body requires continuous erythrocyte replenishment, and it is likely that Epo-EpoR binding in CFU-E progenitors play a central role in the compensatory erythropoiesis during PKD pathology. The significant reduction in the number of CFU progenitors found in the BM of anemic mice, and its increment in the spleen (Figure 25.A), together with the high Epo levels found in plasma, demonstrate the stress erythropoiesis in response to anemia. In addition, differences in erythroid maturation pattern were more evident in spleen (Figure 22), highlighting the role of the spleen in compensatory erythropoiesis response. However, the high levels of 2,3-DPG often reported in patients (Lakomek et al., 1994) (Zanella and Bianchi, 2000) and in the CBA PK-1^{slc}/Pk-1^{slc} mice (Morimoto et al., 1995) suggest that an additional mechanism may be triggering compensatory erythropoiesis in PKD pathology, in addition to Epo-EpoR binding in CFU-Es.

It is thought that specialized stress BFU-E progenitors that reside in the spleen and that are distinct from BM steady state erythropoietic progenitors, are responsible for this stress erythropoiesis mechanism (Paulson et al., 2011). On the other hand, differences found in CFU content in BM and spleen within this PKD model (Figure 25.A) pointed to an active migration of progenitors from the BM to the spleen, where they would adopt the stress erythroid progenitor fate (Perry et al., 2009). This result was also supported by the basal mobilization of CFU progenitors to peripheral blood observed in PKD mice (Figure 25.B PBS injected mice) and observed in some PKD patients (*unpublished data*). However, the CFU mobilization after G-CSF stimulation was clearly reduced in RPK deficient mice, probably as a consequence of the hemolytic state in these mice, leading to an impaired migration of progenitors from the BM to the periphery, and ultimately to the spleen.

The phenotypic characterization studies in this work showed that one of the critical pathological features of PKD is the overexpansion of the erythroid compartment in ectopic hematopoietic sites (spleen and liver) (Figures 23 and 24) as a consequence of the hemolytic anemia, likely promoted by increased Epo levels (controls in Figures 31 and 45). Although increased apoptosis in the erythropoietic progenitors seems the most likely cause of ineffective erythropoiesis, the molecular characterization of erythroid compartment in these PKD mice would offer new insights of PKD pathology. Among the possible candidates to be studied, Jak2 and forkhead box O3 (FOXO3) stand out because of their key role in proliferation, differentiation and survival of erythroid progenitors in response to Epo

(Rivella, 2012). Studies from Marinkovic et al in Foxo3$^{-/-}$ mice have demonstrated the shortened RBC lifespan due to oxidative damage, which displayed increased levels of ROS (Marinkovic et al., 2007). Interestingly, these authors found that FOXO3 loss induced a mitotic arrest only in the population II of erythroid precursors (basophilic erythroblasts), leading to a significant decrease of mature RBC (Marinkovic et al., 2007).

Although until now we have assumed that PKD alterations were limited to the erythroid compartment, the recorded CFU progenitors were considered as total numbers, because in our experiments the hemoglobinization of the colonies was not reliable enough. Nevertheless, the identification of the differently committed colonies in CFUs assays would have allowed us to know whether RPK defects were restricted to the erythroid lineage, or if PKD might also affect the hematopoiesis at an earlier stage of differentiation. This last hypothesis seems unlikely, given the specific nature of the RPK enzyme, but this option has not been considered before. Additionally, differences found in the percentages of LSK and Slam uncommitted progenitors both in the spleen and BM of the PKD anemic mice (Figure 26.B and C) may suggest additional alterations beyond the committed erythroid lineage in the absence of functional RPKs.

Bone marrow from PKD mice showed a reduced population of LSK and Slam progenitors, being more evident in the BM than in the spleen, and more significant when compared to C57BL/6 or AB6F1 mice than with A/J mice (Figure 26.B). In addition, this reduction in the number of pluripotent progenitors did not change the cellularity of the BM, but it markedly increased in the spleen (Figure 26.C). This finding could be explained by the compensatory erythropoiesis mechanism displayed by PKD mice (Figures 22). However, the fact that BM cellularity was unaltered in the presence of lower numbers of LSK and Slam progenitors may suggest an impairment of the pluripotent HSC compartment. Because of the glycolytic defect, the continuously activated stress erythropoiesis in PKD mice turns out to be inefficient, creating a continuous loop that may ultimately stress HSCs located upstream in the hematopoietic hierarchy of differentiation. This idea may suggest that non-functional RPKs could affect other hematopoietic lineages, and in fact anemic mice showed a 30% reduction in the number of WBC when compared with healthy controls (Figure 18.F, Table 7). Although the expression of RPK in leukocytes is not clear yet, it is possible that non-functional RPK may cause subtle alterations in WBC by reducing their metabolic capacity, probably compensated by the expression of other isoforms, such as PKM2 in leukocytes, or

by increasing the cellular metabolism in non-erythroid cells. This idea may be reinforced by the exacerbation of hemolytic anemia observed during infections and pregnancy (Dolan et al., 2002) (Zanella et al., 2007).

Under physiological conditions, primitive HSCs are in a quiescent state in their BM niche, where they preserve their self-renewal capacity while continuously producing all types of blood cells. In response to stress, such as acute anemia, HSCs enter cell cycle and divide to increase blood cells production. Studies from Rossi et al with human HSCs showed that the premature disruption of HSC quiescence exhausted the HSC pool, leading to hematologic failure caused by DNA damage or oxidative, replicative or metabolic stress (Rossi et al., 2007). In addition, other studies have shown that the increase in ROS concentration causes DNA damage in human HSCs, impairing their self-renewal capacity (Yahata et al., 2011). This may explain the reduction of LSK progenitors in PKD anemic mice (Figure 26.A), as well as the reduction of leukocytes found in these animals (Figure 18.F). To our knowledge, hematologic failure has never been reported in PKD patients, but the most severely affected patients fail to survive, and deleterious mutations of *PKLR* gene lead to *in utero* and neonatal death (Gilsanz et al., 1993) (*see Table 2*). Thus, although the possibility of HSC oxidative damage as a consequence of an erythroid enzymatic deficiency seems distant, more detailed studies of PKD have not been conducted so far, and this new possibility opens up a new field to be explored.

Although the reduction of undifferentiated progenitors in PKD deficient mice was clear when compared to any of the control mice used, different genetic backgrounds can account for these differences, as A/J mice displayed percentages of LSK progenitors much lower than C57BL/6 mice (Figure 26). In addition, CFU assays were performed with total BM, so we could not assess whether the functionality of HSC progenitors was affected by RPK deficiency. Even so, the clear differences that we observed in the size of CFU colonies from PKD mice compared with controls (data not shown), and the low ATP concentration in leukocytes from PKD animals that we observed in previous experiments -Table 2 in (Meza et al., 2009)-, encouraged us to perform new experiments towards tracking the RPK role in the different hematopoietic compartments which are currently ongoing.

Recently, it has been seen that the PKM2 isoform displays a role previously unappreciated. It maintains cellular redox homeostasis during metabolic stress through a

process called aerobic glycolysis or the Warburg effect (Hamanaka and Chandel, 2011). Studies from Anastasiou et al have demonstrated that oxidation of PKM2 by increased levels of ROS, fuels pentose phosphate pathway via G6P in cancer cells, producing NADPH and increasing cell oxidative response capacity (Anastasiou et al., 2011). Hence, a glycolytic enzyme becomes a key regulator of cancer progression, emphasizing the importance of glycolysis in general metabolism. This example illustrates that new functions of RPK may emerge, given the key role of PK-catalysed reaction.

Overall these results demonstrated that AcB55 is a suitable model of PKD, reproducing the main symptoms of the human disease: hemolytic anemia, reticulocytosis and splenomegaly. In addition, our phenotypic characterization showed that RPK deficiency affects not only the survival of RBCs, but also the maturation of erythroid progenitors, which results in ineffective erythropoiesis. Unlike most PKD studies, which are focused on the molecular characterization of the different mutations and their association with the phenotype observed in patients, our study covers new features of PKD pathology not studied so far. Thus, AcB55 mice emerge as an ideal strain to assay new therapeutic alternatives for PKD patients.

2. Development of lentivirus-based gene therapy protocols for the treatment of PKD

Currently, there is no specific therapy available for severe cases of PKD, and supportive treatments provide the only options. Allogeneic bone marrow transplantation has been used in specific severe cases, being successful performed in one patient (Tanphaichitr et al., 2000). However, complications derived from allogeneic transplantation make this procedure not available to all patients. Gene therapy via autologous HSC transplantation can be a therapeutic option for severe cases of PKD, also applicable, in theory to all patients, giving its recessive inheritance trait and the restriction of the enzymatic defect to erythroid cells (García-Gómez et al., 2013). Up to now, no gene therapy clinical trials for PKD have been attempted, and preclinical studies are still limited. The low number of patients with poor prognosis requiring BM transplantation, and the lack of

selective advantage of the corrected cells over the diseased ones, have made a gene therapy approach for PKD less attractive than for other monogenic diseases (García-Gómez et al., 2013). Additionally, some RBC pathologies require switching on the expression of the transgene at only the proper stage of differentiation, which represents another challenge in the development of gene therapy protocols. This is the case of RPK, whose expression arises at the stage of basophilic erythroblast and reaches its maximum in mature RBCs.

Despite these hurdles, our previous genetic correction assays in AcB55 PKD mice using γ-RV vectors (Meza et al., 2009) demonstrated that gene therapy would be an effective therapeutic alternative for severe PKD patients lacking an HLA-compatible BM donor. Nevertheless, the recent insights about the oncogenic potential of γ-RV vectors from two hallmark studies (Montini et al., 2009) (Modlich et al., 2009) have made lentiviral vectors replace γ-RV vectors in the gene therapy field. Because of that, we focussed our strategy for PKD treatment on self-inactivating LV vectors as molecular tools for genetic modification. RPK deficient HSCs were transduced with LV carrying a wild type version of the RPK protein, and transplanted into lethally irradiated PKD mice. The efficiency and safety of this protocol in reversing the PKD phenotype were evaluated in the AcB55 deficient mouse model previously characterized. We tested two different ubiquitous promoters, CMV and hPGK, showing that the latter provided efficient levels of phenotypic correction with no observation of transgene silencing events. In addition, we have developed new LV vectors with erythroid-specific expression in order to provide expression-restricted vectors for future PKD gene therapy approaches.

2.1 CMV promoter-based gene therapy protocols

First, we performed experiments based on CMV-carrying LV vectors, as many studies demonstrated a sustained *in vivo* expression of CMV-driven transgenes using LV vectors directed to the liver (VandenDriessche et al., 2002). We optimized the transduction protocol to preserve HSCs self-renewal and functionality (Table 8). According to our results, high MOI (≈ 10 viral particles per cell) and repeated infection cycles (up to three infections) were required for achieving stable and high transduction of HSCs, as no phenotypic correction was achieved in experiments 1, 2 and 4 (Figure 28 and Table 8 and 9) because of either low transduction efficacy or donor engraftment loss (Figure 28.C).

We observed that therapeutic efficacy was highly dependent on donor chimerism and transgene expression levels. With an average of 2 provirus copies per cell and high levels of donor engraftment (90%), the ectopic expression of the therapeutic hRPK transgene was able to rescue PKD hematological pathology (experiment 3, table 8). Deficient mice transplanted with corrected cells showed a steady and long-term correction of the main symptoms reported in PKD patients. As soon as 30 dpt, genetically corrected mice showed an improvement of hematological parameters (RBC, HTC, HGB, MCV and MCH; Figure 29), and a reduction of reticulocytosis (Figure 30) and splenomegaly (Figure 31.A and B). Moreover, normalized levels of circulating Epo (Figure 31.C) decreased the BM compensatory response and the release of reticulocytes to periphery (Figure 30). The reduction of spleen size in animals treated with the therapeutic vector is also an evidence of the reduced compensatory erythropoiesis in response to ameliorated hemolysis.

The requirement of high levels of donor chimerism to achieve therapeutic benefits in this PKD mouse model was confirmed in secondary transplanted receptors, that only showed a partial correction of the PKD phenotype after transplant (Figure 32). This lack of efficient correction was likely caused by the 50% reduction of donor chimerism compared to primary transplanted mice (Table 8). In addition, although blood cells from secondary receptors maintained the number of integrations (2.36 ± 0.66 copies per WBC), the therapeutic hRPK transgene was expressed 2.25 times less compared to primary receptors, suggesting a possible silencing of the CMV promoter (Table 8).

Experiment 5 confirmed the silencing of CMV viral promoter, yielding no expression of the hRPK transgene in WBCs from transplanted mice, in spite of carrying nearly one provirus copy per cell and high levels of donor chimerism (90%). This silencing event corroborated the results obtained by Oertel et al in repopulating experiments with fetal and adult fetal hepatoblasts, transduced with the same lentiviral backbone and promoter expressing the EGFP reporter gene (Oertel et al., 2003), as well as the silencing of the CMV promoter previously reported in γ-RV vectors (Kay, 1992).

The mechanisms that lead to therapeutic gene silencing are classified as dependent or independent of the integration site position, related to the integration of the provirus near or within regions of heterochromatin (Antoniou et al., 2013) and to DNA methylation and histone deacetylation, respectively. Given that LV vectors have a preference towards

integrating in transcriptionally actives units of chromatin (Baum, 2007), and that their genome contains fewer CpG dinucleotides (Pfeifer et al., 2002), it seems likely that the silencing of hRPK transgene in these preliminary assays was caused by DNA methylation of the CMV promoter. This idea is also supported by the CMV susceptibility to CpG methylation reported in adenoviral designs, in contrast to those promoters of housekeeping genes (Brooks et al., 2004).

Transgene gene silencing is a serious problem that can dramatically hamper the therapeutic success of gene therapy protocols, as it has been reported in a clinical trial for X-CGD (Stein et al., 2010). This protocol was based on a γ-RV vector design in which the gp91phox transgene was driven by the SFFV promoter-enhancer element within the vector LTRs (Ott et al., 2006). After initial gene therapy success in two X-CGD patients, the transgene became silenced due to methylation of the SFFV promoter, no longer providing a therapeutic effect (Stein et al., 2010). These outcomes highlight the importance of vector design for developing efficient and safe gene therapy protocols with therapeutic applications.

2.2 Human PGK-based gene therapy protocols

To circumvent the silencing observed in our first set of experiments, we modified the vector design replacing the CMV viral promoter with the human PGK eukaryotic promoter because, although it is a much weaker promoter than CMV, it drives a more stable and sustainable expression of the transgenes (Naldini, 2011) (Gerolami et al., 2000). In addition, as CMV-based assays showed a correlation between high transgene expression and phenotypic correction, we designed and introduced in the vector an optimized version of the human cDNA PKLR gene (coRPK) to increase transcript stability (Figure 12) (*see section 11.1.2 of Materials and Methods*).

The coRPK sequence was designed with increased GC content to improve transgene transcription, although the use of GC-rich sequences of transgenes is controversial, since it could otherwise lead to therapeutic gene silencing. Some studies have shown an increase in expression levels when CpG depleted cassettes were used, in both transgenic mice (Chevalier-Mariette et al., 2003) and somatic cells (Yew et al., 2002). However, the presence of CpG-rich regions in the transgene might otherwise be beneficial for posttranscriptional

processes, as they significantly increased RNA stability (Duan and Antezana, 2003). This idea has been recently demonstrated, reporting a clear loss of EGFP expression at a transcriptional level when CG-depleted transgenes were used, together with a significant decrease of steady-state RNA copy numbers (Bauer et al., 2010). In addition, the efficacy of codon-optimized designs to increase therapeutic transgene expression has been demonstrated in several preclinical protocols. For instance, gene therapy approaches for RAG2-SCID immunodeficiency (van Til et al., 2012), and X1-SCID (Huston et al., 2011) in which the optimized *IL2RG* transgene yielded higher expression than its non-optimized counterpart, both with a similar number of integrations per cell (Huston et al., 2011).

The newly developed LV vector was used to perform two independent experiments with a relatively high number of transplanted animals (Figure 34 and table 13), demonstrating that hPGK promoter was potent enough to express clinically relevant levels of RPK protein. Our results address for the first time an efficient phenotypic correction of PKD disease through LV-based gene therapy, likely as a consequence of sufficient levels of coRPK transgene when was driven by the hPGK promoter.

Bone marrow progenitors from PKD anemic mice were transduced with the PGK-lentiviral vector and transplanted into lethally irradiated deficient mice. We achieved a long-term PKD correction in coRPK expressing mice (up to 9 months post-transplantation), including the complete reversion of hematopoietic variables (RBC, HGB, HTC, MCV and MCH) (Figure 35 and table 10) that reached values close to C57BL/6 wild type mice. Constitutive reticulocytosis, a typical sign of PKD disease, was significantly reduced in genetically corrected mice (Figure 36), showing a marked improvement as soon as 30 dpt compared to PKD non-transplanted mice or mice transplanted with EGFP-carrying LV vector. However, genetically corrected animals showed 10% residual levels of reticulocytes along the experiment, probably caused by the 36.34% remaining endogenous hematopoiesis detected in transplanted mice (Table 11).

RBC half-life also improved in transplanted mice as a consequence of the ectopic expression of the human RPK protein, extending up to 6 days compared to non-transplanted controls (Figures 37 and 38). However, RBC half-life data obtained in the gene therapy experiments differ from those obtained during the phenotypic characterization for all groups of mice, as well as from previously reported data (Min-Oo et al., 2007). We think these discrepancies are caused by different efficiencies of the *in vivo* RBC labelling, since we

followed the same experimental procedure in all experiments. In addition, C57BL/6 healthy controls unexpectedly displayed labelled-RBC half-lives similar to anemic control mice (Figure 38.B), although their kinetic of RBC survival was always above both transplanted and non-transplanted PKD mice (Figure 38.A).

Nonetheless, genetically corrected mice showed a significantly extended RBC half-life compared to deficient PKD mice, displaying an intermediate RBC survival phenotype. This partial RBC survival rescue agrees with 63.66% donor chimerism detected in these mice, which carried 1.65 provirus copies per blood cell (Table 11). Interestingly, although the initial labelled-RBC percentage in RPK expressing mice was much lower than in healthy controls (≈24% and ≈87%, respectively), this percentage rose up to ≈54% as soon as 2 days after labelling (Figure 38.A). This enrichment of labelled-RBC may be explained by an increased half-life of corrected RBCs over the deficient ones, suggesting a survival advantage of corrected mature erythrocytes expressing the therapeutic transgene in circulation. Our results demonstrate that LV-derived functional RPK proteins were able to extend RBC survival in PKD pathology, likely conferring a competent metabolic status and decreasing potential erythrocyte membrane alterations.

Reduced hemolytic anemia in RPK-transplanted mice as a consequence of extended RBC half-life ameliorated the reticulocyte release response from the bone marrow, and it was also reflected in a significant reduction of splenomegaly (Figure 39), probably as a consequence of the reduced spleen activity in removing damaged erythrocytes from circulation. The genetic modification led to a complete reversion of splenomegaly compared to PKD controls or PKD animals transplanted with EGFP-carrying vector, even with a moderate percentage of transduced cells (65% of CFU progenitors carrying the therapeutic transgene) (Table 11). Unexpectedly, control mice expressing the EGFP transgene showed a reduction in spleen weight compared to PKD anemic mice (Figure 39), which might be related to the proliferative stage of progenitors after the transplant.

Moreover, the improvement of the hemolytic status in PKD mice after gene therapy led to a clear reversion of organ pathology. Mice carrying the ectopic coRPK transgene showed normal histological structure in the spleen (Figure 40) and liver (Figure 41), displaying RBC accumulations less frequently than PKD mice or anemic mice expressing the EGFP control transgene. Remarkably, the expression of the human RPK carried by the vector

was enough to rescue the iron overload-derived complications in the liver (Figure 41), highlighting the therapeutic potential of this gene therapy LV vector. These results demonstrated a reduction of extramedullar erythropoiesis in mice transplanted with coRPK-transduced progenitors, which was supported by the normalization of erythroid maturation pattern in both BM and spleen (Figure 42), and the significantly reduced levels of Epo in the plasma of these animals (Figure 45).

Through the study of the different erythroid subpopulations, we demonstrated that LV-derived RPK expression in erythroid cells was needed to correct erythrocyte maturation and generate normal RBCs (Figure 42). The increment of immature erythroid precursors (I and II) that PKD mice showed to compensate the quick loss of deficient RBCs, is not enough to increase the number of mature RBCs (IV) (Figure 43). However, the genetic correction in transplanted mice was able to stabilize the erythroid maturation to normal levels, rescuing the ineffective erythropoiesis typical from PKD mice. Mice carrying the therapeutic transgene showed a reduction of proerythroblast and basophilic erythroblast (I and II) in parallel with healthy controls, as well as a significant increase of the latest erythroid compartment (IV) (Figure 43). Thus, ectopic coRPK expression leads to an improved metabolic status of the basophilic erythroblasts, stage in which RPK expression starts to be predominant, allowing the appropriate differentiation into the next stage of erythroid maturation.

Overall, the reduced RBC hemolysis and extended RBC half-life as a consequence of coRPK transgene expression completely reverted the compensatory erythropoiesis in transplanted mice from the therapeutic group, no longer needing such as intense RBC replenishment. As we commented in the first part of the discussion, aberrant RPKs lead to inefficient compensatory erythropoiesis likely caused by an increased apoptosis in the erythroid compartment, suggesting a survival advantage of corrected cells in the erythroid committed compartment. This hypothesis has been demonstrated in the transgenic rescue of CBA PK-1^{slc}/Pk-1^{slc} deficient mice, which showed reduced apoptosis in the spleen correlated with the expression level of wild type RPK protein (Kanno et al., 2007). The expression of the therapeutic transgene acted by reverting the compensatory mechanisms, as erythropoiesis became efficient in the presence of functional RPK enzymes. In addition, the decreased Epo levels found in the plasma of genetically corrected mice (Figure 45) yielded a significant reduction of CFU progenitors in the spleen, reaching numbers close to

C57BL/6 controls (Figure 44.B). However, CFU content in bone marrow was unaltered after the genetic modification compared to PKD anemic mice or EGFP controls (Figure 44.A). These results confirm that mouse stress erythropoiesis is mainly located in the spleen, explaining the lack of changes in the bone marrow CFU content.

The stability and safety of the gene therapy protocol developed in this work was also demonstrated in secondary transplanted mice, showing once more that both viral dosage and donor engraftment were relevant parameters for achieving therapeutic success in PKD gene therapy. The genetic modification of bone marrow cells transplanted into secondary receptors led to an efficient phenotypic correction in experiment 2.2, with just 62.89% of donor chimerism and 1.44 provirus copies per blood cell (Table 12). These data were comparable to those obtained in primary transplanted mice, as well as the level of correction achieved in the hematological parameters (Figures 48-50) and spleen pathology (Figure 50). In addition, the hematopoietic reconstitution was multi-lineage, confirming the transplantation of functional pluripotent HSCs with no imbalance in the different hematopoietic lineages (Figure 52).

Unexpectedly, some of the secondary receptors from experiment 2.1 developed leukaemia 60 days after transplant. We think this event was likely caused by the high inbreeding degree in our PKD colony, as leukemic mice belonged to both groups of transplanted animals: coRPK and EGFP expressing mice (Appendix 2). Therefore, a metabolic advantage caused by RPK overexpression was discarded as the cause of leukaemia onset. On the other hand, no proviral integrations were detected in blood or spleen cells from leukemic mice, also EGFP expression was not detected in leukemic mice from the control group (Figure 47). Therefore, we concluded that this was an isolated event (8 out of 88 transplanted surviving mice) and that lentiviral integration was not the cause of insertional oncogenesis in this set of secondary transplanted mice. Despite this, the integration pattern of both lentiviral vectors is being currently analyzed, and tertiary transplants are indeed ongoing with no signs of leukaemia three months post-transplantation (data not shown), demonstrating the stability and safety of this gene therapy treatment.

As we mentioned before, levels of donor engraftment were determinant to achieve clinical benefits in this PKD mouse model, because a selective advantage of corrected HSC has not been reported for this disease. This implies the need of myeloablative conditioning

to favour the repopulation of bone marrow niches by transduced HSCs, as was demonstrated by Zaucha et al in Basenji dog PKD model (Zaucha et al., 2001). However, previous allogeneic transplantation studies in this model showed that approximately 20% of corrected long-term repopulating cells are required to ameliorate the disease phenotype after non-myeloablative conditioning (Takatu et al., 2003).

A different gene therapy approach using the CBA PK-1^{slc}/Pk-1^{slc} deficient mouse model showed that anemia could be corrected when 10% of engrafted cells expressed the endogenous *Pklr* gene (Richard et al., 2004), and our previous gene therapy assays reported similar results (Meza et al., 2009). Using a γ-RV vector and different proportions of corrected and non-corrected cells transplanted into AcB55 deficient mice, we demonstrated that long-term phenotypic correction of PKD was achieved with a minimum of 25% of corrected repopulating cells (Meza et al., 2009). However, the lack of healthy donor mice with a compatible genetic background did not allow us to perform competition transplants in this PKD model, required to determine the minimum number of cells corrected by LV-transfer to achieve an efficient phenotypic correction. Nevertheless we observed a more efficient correction in primary transplanted mice from assay 1 (Table 11) than in the assay 2, likely due to a higher transduction percentage of CFU progenitors in genetically corrected mice (91.08%). The higher number of provirus copies also accounted for these differences (Table 11), as a higher number of copies would lead to an increase in expression of human monomers, increasing the probability of building up functional RPK tetramers. However, pathological manifestations only appear in PKD patients when RPK enzymatic activity falls below 25% of the normal activity (Zanella et al., 2001a), so it is reasonable to think that a therapeutic benefit would be obtained from PKD gene therapy even if normal levels of protein expression are not achieved.

Metabolomic studies also revealed interesting insights about the metabolic rescue of PKD via gene therapy. The untargeted metabolic profiling study showed two clusters of genes in which the trend of metabolite expression in RBCs changed in response to the ectopic expression of the RPK (Figure 53). Glycolytic metabolite profile expression was more similar between genetically corrected mice and healthy controls, than to animals with the same genetic background (non-transplanted or EGFP-transplanted PKD mice) for cluster 1 and 2 (Figure 53), indicating a correction of the glycolytic pathway in an expression level. In the third cluster, however, similarities were not as clear, and the metabolite array showed

some metabolites whose expression was slightly increased (Figure 53 assay 2). Metabolite profile from cluster 3 in genetically corrected cells was therefore more similar to PKD mice carrying the EGFP control transgene than to C57BL/6 healthy controls, which may be related to the different genetic background or to the lentiviral integration itself. Repeated metabolic profiling with a higher number of samples showed again that gene therapy treatment based on the coRPK-carrying LV vectors was able to correct PKD phenotype in terms of glycolytic metabolite content. The expression of the therapeutic transgene modified the profile of metabolites involved in glycolytic processes, reaching a pattern similar to non-deficient mice for the selected metabolites (Figure 54). In addition, targeted LC/MS profiling showed an increased of ATP and ADP in genetically corrected RBCs, pointing to an improved energetic status of these cells (data not shown). On the other hand, the reduced expression trend in metabolites located upstream of the RPK-catalysed reaction, such as 3-phosphoglycerate and PEP (Figure 55), the direct substrate of RPK enzyme, demonstrate the recovery of glycolytic pathway functionality.

Besides, the same analysis performed in parallel on WBC from the same animals, revealed that ectopic coRPK expression rescued the glycolytic pathway in RBCs but did not change the metabolic balance in WBCs (Figure 56). This finding is greatly important in terms of gene therapy safety, as the ubiquitous expression of coRPK transgene driven by the hPGK promoter, could potentially disturb ATP production in WBC providing them with a metabolic advantage. Nonetheless, although the number of WBCs did not change after the gene therapy treatment (Table 10), the lack of differences in the metabolite expression profile of WBCs (Figure 56) confirms the absence of non-desired side effects in leukocytes.

The lentiviral vector developed in this work provides several advantages over the previously tested ones. First, the use of a SIN-LV vector design allowed a relatively easy and safe production of viral stocks with high titer, able to efficiently transduce RPK-deficient mouse HSCs. Vector sequence also include several modifications to improve transgene expression in target cells. The use of a weak promoter such as hPGK, which is less susceptible to silencing by methylation (Gerolami et al., 2000), leads to a more physiological expression of the transgene, achieving therapeutic levels with a viral dosage within the clinical standards (Matrai et al., 2010). In addition, the transgene codon optimized sequence design increases transgene mRNA stability, and no reporter gene was included in the therapeutic vector sequence, avoiding possible immunogenic problems (Morris et al., 2004)

(Stripecke et al., 1999). Overall, these preclinical results pave the way towards the development of future clinical trials for severely affected PKD patients lacking HLA-compatible donor.

Altogether, these data demonstrate that the developed hPGK-coRPK LV vector efficiently reverted PKD pathology in both primary and secondary deficient mice transplanted with progenitors carrying 1.65 copies per cell of the therapeutic transgene. Human PGK promoter was potent enough to express clinically relevant levels of RPK protein when donor chimerism was ≥ 60%, restoring the hemolytic phenotype in transplanted mice. This genetic correction was not only able to extend RBC half-life, normalizing the hematological variables and reticulocytes levels, but also to revert the compensatory erythropoiesis constitutively activated in PKD mice, rescuing the pathology in the organs affected by iron overload complications. These findings demonstrate that hPGK promoter drove stable and robust expression of the coRPK transgene, suggesting that hPGK is able to resist the epigenetic changes that proerythroblasts undergo upon differentiation into mature RBCs. In addition, the ectopic expression of human RPK corrected the energetic defect in RBCs without altering the metabolic balance in WBCs, underscoring the efficacy and safety of this gene therapy protocol.

3. Development of erythroid-specific lentiviral vectors for the PKD

As previously mentioned, one of the major drawbacks in developing genetic correction treatments for PKD disease is the lack of selective advantage of corrected HSCs over deficient HSCs. However, this could otherwise be an advantage as, theoretically, genetic correction is only required in mature RBCs. Red blood cells have a low proliferative capacity and short half-life compared to other cell types and they become transcriptionally inactive upon differentiation, decreasing the likelihood of insertional oncogenesis. In addition, the development of vectors with lineage-restricted expression of the transgene reduces potential adverse effects in non-affected lineages. On the other hand, the use of βLCR sequences in lentiviral vectors offers additional benefits for developing therapeutic vectors including their ability to enhance the expression of linked genes to physiological

levels in a tissue-specific and copy number-dependent manner (Li et al., 2002). Also, recent studies have shown that LCR enhanced-lentiviral vectors have a lower propensity to disturb host gene function, and thus a reduced insertional mutagenesis potential (Emery, 2011); (Ramezani and Hawley, 2010).

With this aim, we developed several erythroid-tissue specific vectors based on SIN HIV1-derived backbones. The first design included the regulatory sequences of the gene to be corrected (PKLR promoter) (Figure 14) and the coRPK transgene (Figure 15). This physiologically regulated lentiviral vector may have significant advantages for clinical applications in humans, as transgene expression is controlled by natural gene regulation (Toscano et al., 2011). However, although driving the expression of the therapeutic gene using its own promoter seems the most logical approach, our preliminary experiments showed that PKLR-driven expression was notably low (data not shown). Therefore, we included the regulatory sequences of β-LCR upstream of PKLR promoter to foster both specificity and potency of the PKLR promoter (Figure 16). The usefulness of LCR combinations has been demonstrated in previous works, rendering therapeutic levels of β-globin expression with a viral dose of 1-2 provirus copies per cell (May et al., 2000); (Puthenveetil et al., 2004). In addition, we developed another LV vector carrying the HS2/HS3 enhancers and the β-globin promoter as a control erythroid promoter, driving the expression of coRPK transgene (Figure 16). This combination has been previously tested in a mouse model of β-thalassemia (Miccio et al., 2008).

One of the most important features of β-LCR is its strong transcription-enhancing activity (Fraser and Grosveld, 1998) (Navas et al., 2001), so we studied the degree of specificity of these three new coRPK-carrying vectors by transduction of erythroid (K562 and HEL) and non-erythroid (HL60) human cell lines. Our results demonstrate that human PKLR promoter drove an erythroid-specific expression of the transgene, detecting the either EGFP (Figure 57) or coRPK transgene expression (Figure 58) in only the erythroid cell lines (K562 and HEL), while the human PGK promoter led to ubiquitous expression of the transgenes (Figure 57 and 58) at similar levels of expression among the different cell lines (Figure 59). On the other hand, K562 was the only cell line that expressed the endogenous RPK protein. However, because no differentiation protocols were applied, HEL cells are expected to increase endogenous RPK expression upon differentiation.

Quantification of the ectopic RPK expression allowed us to compare the level of expression conferred by each combination of erythroid promoters and βLCR sequences. According to our results, HS3-HS2-β globin promoter was the combination that provided the highest expression of coRPK transgene, being 1.8-2.3-fold higher than with the PKLR promoter in the erythroid cell lines (Figure 59), and even higher than the coRPK expression driven by hPGK promoter. However, vector carrying the PKLR promoter alone led to a higher coRPK expression in HEL cells than in K562, which suggests a more physiological expression of the transgene. Unexpectedly, βLCR sequences did not enhance coRPK expression when PKLR promoter was used (Figure 59), which might be related with the size of the vector (12.7 kb, Figure 16). Although all experiments were carried out in parallel with the same number of cells and the same MOI, further studies will be required to address the potency of these erythroid enhancers/promoter combinations driving the coRPK expression.

Additionally, we are currently developing a different erythroid-specific transgene expression strategy based on microRNAs (miRNA) (Brown et al., 2007b). The miR223 has been proposed as a myeloid-specific regulator that negatively controls progenitor proliferation and granulocyte differentiation and activation (Johnnidis et al., 2008). Studies by Felli et al have shown that the down-regulation of miR223 was required for erythroid differentiation, leading to expansion of the erythroblast population (Felli et al., 2009). The incorporation of the miR223 target sequences in the 3' UTR of the transgene would specifically supress coRPK in the cells that express miR223, leading to a de-targeted coRPK expression in unwanted hematopoietic lineages. In addition, because miRNA control is post-transcriptional, miRNA223 target strategy could be combined with other transcriptional elements to obtain more specific transgene expression patterns in future vectors (Brown and Naldini, 2009).

Nevertheless, the preclinical evaluation of the developed erythroid-specific vectors to be used in future gene therapy trials will require the study of their functionality in PKD patient $CD34^+$ cells. PKD disease has, nonetheless, an additional handicap, as only peripheral blood samples are used for its diagnosis and bone marrow harvest is not recommended. This limits the number of PKD patient $CD34^+$ cells that could be used for testing gene therapy vectors. Because of that, we are currently developing a silencing strategy based on RNA interference (Davidson and McCray, 2011) to inhibit the endogenous RPK expression in wild type $CD34^+$ cells. The 80.4% homology between coRPK transgene and the human wild type *PKLR* gene (Appendix 1) allowed the design of several siRNA that specifically recognize the

endogenous *PKLR* gene sequence, but not the optimized one. This work will hopefully provide an artificial model of human PKD disease, enabling the study of coRPK expression upon erythroid-specific LV vectors transduction to test its therapeutic potential in RPK deficient primary cells.

4. Final remarks

The lack of satisfactory therapeutic options for severe PKD patients, together with a better understanding of the physiopathological mechanisms involved in the disease and reported experimental evidence of genetic correction, provide an ideal scenario for developing gene therapy approaches for PKD. Considering the occurrence of unexpected adverse events due to insertional mutagenesis in clinical trials using γ-RV vectors, we focused our gene therapy strategy on a self-inactivating lentiviral vector harbouring the weak and physiological hPGK promoter. This design allowed the long-term correction of mouse PKD disease with no adverse effects, achieving high levels of donor engraftment without any selection method of genetically corrected cells. Our preclinical results provide encouraging expectations that HSC genetic correction with such a vector could be a therapeutic option for severely affected PKD patients lacking HLA-compatible donor, paving the way towards the development of future clinical trials.

Currently, gene replacement addition strategy appears to be the most reliable applicable strategy. The gene addition protocol proposed in this work would be able to correct the wide range of mutations described in PKD patients, including mutations in the regulatory sequence of the gene. On the other hand, as *PKLR* null mutations are rare, most patients display truncated versions of the RPK protein or non-functional complete proteins, avoiding a possible immunological response against the exogenous protein in gene therapy-transplanted patients.

Nevertheless, the ubiquitous nature of hPGK promoter made us consider a new genetic correction strategy based on enhanced tissue-specific or physiological-regulated vectors. The targeted coRPK expression to the erythroid lineage may lead to a decreased risk of adverse effects in other cell-types not affected by the disease, such as leukocytes or

granulocytes. However, as we speculated at the beginning of this section, the RPK enzyme may have additional functions beyond the erythroid compartment. Although we are currently exploring this possibility, the demonstration of this idea would support the use of the hPGK ubiquitous promoter in gene therapy vectors for PKD, providing the best balance between efficacy and safety for future clinical trials.

VIII. RESULTS SUMMARY

1. AcB55 mice reproduced the hematological symptoms of human PKD disease, showing a significant reduction of RBC counts, hemoglobin levels and hematocrit index, together with acute splenomegaly and constitutive reticulocytosis. In addition, AcB55 mice show an intense liver iron overload, reproducing one of the major problems of severely affected PKD patients.

2. The 269 T>A loss-of-function mutation in the *Pklr* gene led to shortened RBC half-life, triggering compensatory mechanisms (extramedullar or stress erythropoiesis and reticulocytosis) involving high levels of Epo to ameliorate the hemolytic anemia. RPK deficiency affected not only the survival of RBCs, but also the maturation of erythroid progenitors in bone marrow and spleen, which resulted in ineffective erythropoiesis failing to increase the production of mature RBCs.

3. PKD mice show a reduction in the number of pluripotent progenitors (LSK and Slam progenitors) both in the bone marrow and spleen, and total CFU progenitors were also reduced in the bone marrow. However, in the spleen CFU content was increased, pointing to an active extramedullar erythropoiesis. In addition, RPK deficiency led to an impaired migration of progenitors from the BM to the periphery upon G-CSF stimulation.

4. Gene therapy approach with the pRRL-CMV-hRPK developed lentiviral vectors provided a partial phenotypic correction of transplanted PKD mice because of low transduction efficacy and CMV promoter silencing.

5. Gene therapy protocol based on the developed pCCL-hPGK-coRPK lentiviral vector efficiently reverted PKD pathology in both primary and secondary deficient transplanted mice, with an average dosage of 1.65 integrated copies per cell. Human PGK promoter was strong enough to express clinically relevant levels of human RPK protein when donor chimerism was ≥ 60%, normalizing the hematological variables and reticulocyte levels, extending RBC half-life and reverting the compensatory erythropoiesis.

6. The hemolytic anemia reversion after hPGK-coRPK genetic correction rendered normal levels of Epo, decreasing the number of spleen CFU progenitors to control

values and rescuing the pathology in the organs, including reversion of extramedullar erythropoiesis and iron overload.

7. The ectopic expression of human RPK corrected the energy deficit in mouse RBCs without altering the metabolic balance in WBC, underscoring the efficacy and safety of this gene therapy protocol.

8. Lentiviral vectors carrying either the β globin or the PKLR erythroid promoter were highly specific in transduction *in vitro* studies. The HS3-HS2-β globin promoter was the combination that provided the highest expression of coRPK transgene, but PKLR promoter showed a high erythroid-specificity even in the absence of βLCR enhancer sequences. In addition, ectopic coRPK expression was higher in HEL cells than in K562, suggesting a more physiological expression of the coRPK transgene when was driven by its own promoter.

IX. CONCLUSIONS

1. AcB55 mice carrying the 269 T>A loss-of-function mutation in the *Pklr* gene are a suitable model of human pyruvate kinase deficiency, reproducing the main symptoms observed in PKD patients, allowing the study of the pathological mechanisms involved in PKD pathology and providing an ideal scenario for developing gene therapy strategies.

2. Developed hPGK-coRPK self-inactivating lentiviral vector carrying the human PGK ubiquitous promoter and the codon optimized sequence of the human *PKLR* gene cDNA, efficiently revert the hemolytic phenotype in AcB55 mouse model through transduction and transplantation of deficient hematopoietic progenitors, emerges as an efficient and safe gene therapy vector for future clinical trials focussed on severe PKD patients lacking HLA-compatible donors.

3. We have provided new lentiviral vectors carrying the erythroid tissue-specific promoters from the β-globin and *PKLR* genes that were highly specific in *in vitro* transduction studies. PKLR promoter drove a highly erythroid-tissue specific expression of the coRPK transgene in human erythroid cell lines, even in the absence of βLCR enhancers, providing a new physiologically-regulated vector that may be useful for future PKD gene therapy approaches.

X. BIBLIOGRAPHY

1. Abu-Melha, A.M., Ahmed, M.A., Knox-Macaulay, H., Al-Sowayan, S.A., and el-Yahia, A. (1991). Erythrocyte pyruvate kinase deficiency in newborns of eastern Saudi Arabia. Acta Haematol *85*, 192-194

2. Aisaki, K., Aizawa, S., Fujii, H., Kanno, J., and Kanno, H. (2007). Glycolytic inhibition by mutation of pyruvate kinase gene increases oxidative stress and causes apoptosis of a pyruvate kinase deficient cell line. Exp Hematol *35*, 1190-1200

3. Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., Scaramuzza, S., Andolfi, G., Mirolo, M., Brigida, I.*, et al.* (2009). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med *360*, 447-458

4. Aiuti, A., Slavin, S., Aker, M., Ficara, F., Deola, S., Mortellaro, A., Morecki, S., Andolfi, G., Tabucchi, A., Carlucci, F.*, et al.* (2002). Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science *296*, 2410-2413

5. Aizawa, S., Harada, T., Kanbe, E., Tsuboi, I., Aisaki, K., Fujii, H., and Kanno, H. (2005). Ineffective erythropoiesis in mutant mice with deficient pyruvate kinase activity. Exp Hematol *33*, 1292-1298

6. Aizawa, S., Kohdera, U., Hiramoto, M., Kawakami, Y., Aisaki, K., Kobayashi, Y., Miwa, S., Fujii, H., and Kanno, H. (2003). Ineffective erythropoiesis in the spleen of a patient with pyruvate kinase deficiency. Am J Hematol *74*, 68-72

7. Almarza, D., Bussadori, G., Navarro, M., Mavilio, F., Larcher, F., and Murillas, R. (2011). Risk assessment in skin gene therapy: viral-cellular fusion transcripts generated by proviral transcriptional read-through in keratinocytes transduced with self-inactivating lentiviral vectors. Gene Ther 18, 674-681

8. Alonso-Ferrero, M.E., Valeri, A., Yanez, R., Navarro, S., Garin, M.I., Ramirez, J.C., Bueren, J.A., and Segovia, J.C. (2011). Immunoresponse against the transgene limits hematopoietic engraftment of mice transplanted in utero with virally transduced fetal liver. Gene Ther 18, 469-478

9. Anastasiou, D., Poulogiannis, G., Asara, J.M., Boxer, M.B., Jiang, J.K., Shen, M., Bellinger, G., Sasaki, A.T., Locasale, J.W., Auld, D.S., et al. (2011). Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. Science 334, 1278-1283

10. Antoniou, M., Harland, L., Mustoe, T., Williams, S., Holdstock, J., Yague, E., Mulcahy, T., Griffiths, M., Edwards, S., Ioannou, P.A.*, et al.* (2003). Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. Genomics *82*, 269-279

11. Antoniou, M.N., Skipper, K.A., and Anakok, O. (2013). Optimizing retroviral gene expression for effective therapies. Hum Gene Ther 24, 363-374

12. Antoniou, M., and Grosveld, F. (1990). beta-globin dominant control region interacts differently with distal and proximal promoter elements. Genes Dev *4*, 1007-1013

13. Arumugam, P.I., Higashimoto, T., Urbinati, F., Modlich, U., Nestheide, S., Xia, P., Fox, C., Corsinotti, A., Baum, C., and Malik, P. (2009a). Genotoxic potential of lineage-specific lentivirus vectors carrying the beta-globin locus control region. Mol Ther *17*, 1929-1937

14. Arumugam, P.I., Scholes, J., Perelman, N., Xia, P., Yee, J.K., and Malik, P. (2007). Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element. Mol Ther *15*, 1863-1871

15. Arumugam, P.I., Urbinati, F., Velu, C.S., Higashimoto, T., Grimes, H.L., and Malik, P. (2009b). The 3' region of the chicken hypersensitive site-4 insulator has properties similar to its core and is required for full insulator activity. PLoS One *4*, e6995

16. Bank, A., Dorazio, R., and Leboulch, P. (2005). A phase I/II clinical trial of beta-globin gene therapy for beta-thalassemia. Ann N Y Acad Sci *1054*, 308-316

17. Baronciani, L., and Beutler, E. (1995). Molecular study of pyruvate kinase deficient patients with hereditary nonspherocytic hemolytic anemia. J Clin Invest *95*, 1702-1709

18. Baronciani, L., and Beutler, E. (1993). Analysis of pyruvate kinase-deficiency mutations that produce nonspherocytic hemolytic anemia. Proc Natl Acad Sci U S A *90*, 4324-4327

19. Baronciani, L., and Beutler, E. (1994). Prenatal diagnosis of pyruvate kinase deficiency. Blood 84, 2354-2356

20. Bauer, A.P., Leikam, D., Krinner, S., Notka, F., Ludwig, C., Langst, G., and Wagner, R. (2010). The impact of intragenic CpG content on gene expression. Nucleic Acids Res 38, 3891-3908

21. Baum, C. (2007). Insertional mutagenesis in gene therapy and stem cell biology. Curr Opin Hematol *14*, 337-342

22. Baum, C., Schambach, A., Bohne, J., and Galla, M. (2006). Retrovirus vectors: toward the plentivirus? Mol Ther *13*, 1050-1063

23. Berghout, J., Higgins, S., Loucoubar, C., Sakuntabhai, A., Kain, K.C., and Gros, P. (2012). Genetic diversity in human erythrocyte pyruvate kinase. Genes Immun *13*, 98-102

24. Beutler, E., Dyment, P.G., and Matsumoto, F. (1978). Hereditary nonspherocytic hemolytic anemia and hexokinase deficiency. Blood *51*, 935-940

25. Beutler, E., and Gelbart, T. (2000). Estimating the prevalence of pyruvate kinase deficiency from the gene frequency in the general white population. Blood *95*, 3585-3588

26. Beutler, E., Gelbart, T., and Pegelow, C. (1986). Erythrocyte glutathione synthetase deficiency leads not only to glutathione but also to glutathione-S-transferase deficiency. J Clin Invest *77*, 38-41

27. Biffi, A., Capotondo, A., Fasano, S., del Carro, U., Marchesini, S., Azuma, H., Malaguti, M.C., Amadio, S., Brambilla, R., Grompe, M.*, et al.* (2006). Gene therapy of metachromatic leukodystrophy reverses neurological damage and deficits in mice. J Clin Invest *116*, 3070-3082

28. Biffi, A., and Naldini, L. (2007). Novel candidate disease for gene therapy: metachromatic leukodystrophy. Expert Opin Biol Ther *7*, 1193-1205

29. Black, J.A., Rittenberg, M.B., Standerfer, R.J., and Peterson, J.S. (1978). Hereditary persistence of fetal erythrocyte pyruvate kinase in the Basenji dog. Prog Clin Biol Res *21*, 275-295

30. Blom van Assendelft, G., Hanscombe, O., Grosveld, F., and Greaves, D.R. (1989). The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. Cell *56*, 969-977

31. Boas, F.E., Forman, L., and Beutler, E. (1998). Phosphatidylserine exposure and red cell viability in red cell aging and in hemolytic anemia. Proc Natl Acad Sci U S A *95*, 3077-3081

32. Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P.P., Diez, I.A., Dewey, R.A., Bohm, M., Nowrouzi, A., Ball, C.R., Glimm, H*., et al.* (2010). Stem-cell gene therapy for the Wiskott-Aldrich syndrome. N Engl J Med *363*, 1918-1927

33. Breda, L., Casu, C., Gardenghi, S., Bianchi, N., Cartegni, L., Narla, M., Yazdanbakhsh, K., Musso, M., Manwani, D., Little, J*., et al.* (2012). Therapeutic hemoglobin levels after gene transfer in beta-thalassemia mice and in hematopoietic cells of beta-thalassemia and sickle cells disease patients. PLoS One *7*, e32345

34. Brooks, A.R., Harkins, R.N., Wang, P., Qian, H.S., Liu, P., and Rubanyi, G.M. (2004). Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. J Gene Med 6, 395-404

35. Brown, B.D., Cantore, A., Annoni, A., Sergi, L.S., Lombardo, A., Della Valle, P., D'Angelo, A., and Naldini, L. (2007a). A microRNA-regulated lentiviral vector mediates stable correction of hemophilia B mice. Blood *110*, 4144-4152

36. Brown, B.D., Gentner, B., Cantore, A., Colleoni, S., Amendola, M., Zingale, A., Baccarini, A., Lazzari, G., Galli, C., and Naldini, L. (2007b). Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. Nat Biotechnol *25*, 1457-1467

37. Brown, B.D., and Naldini, L. (2009). Exploiting and antagonizing microRNA regulation for therapeutic and experimental applications. Nat Rev Genet 10, 578-585

38. Brown, B.D., Sitia, G., Annoni, A., Hauben, E., Sergi, L.S., Zingale, A., Roncarolo, M.G., Guidotti, L.G., and Naldini, L. (2007c). In vivo administration of lentiviral vectors triggers a type I interferon response that restricts hepatocyte gene transfer and promotes vector clearance. Blood 109, 2797-2805

39. Bulfield, G., Moore, E.A., and Kacser, H. (1978). Genetic variation in activity of the enzymes of glycolysis and gluconeogenesis between inbred strains of mice. Genetics 89, 551-561

40. Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S., and Hoffmann, C. (2005). Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol 3, 848-858

41. Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Bougneres, P., Schmidt, M., Kalle, C.V., Fischer, A., Cavazzana-Calvo, M., and Aubourg, P. (2012). Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. Methods Enzymol 507, 187-198

42. Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Veres, G., Schmidt, M., Kutschera, I., Vidaud, M., Abel, U., Dal-Cortivo, L., Caccavelli, L*., et al.* (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. Science *326*, 818-823

43. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K*., et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature *467*, 318-322

44. Chang, Y.J., and Huang, X.J. (2013). Donor lymphocyte infusions for relapse after allogeneic transplantation: when, if and for whom? Blood Rev 27, 55-62

45. Chang, K.H., Tam, M., and Stevenson, M.M. (2004). Inappropriately low reticulocytosis in severe malarial anemia correlates with suppression in the development of late erythroid precursors. Blood 103, 3727-3735

46. Charrier, S., Dupre, L., Scaramuzza, S., Jeanson-Leh, L., Blundell, M.P., Danos, O., Cattaneo, F., Aiuti, A., Eckenberg, R., Thrasher, A.J*., et al.* (2007). Lentiviral vectors targeting WASp

expression to hematopoietic cells, efficiently transduce and correct cells from WAS patients. Gene Ther *14*, 415-428

47. Charrier, S., Ferrand, M., Zerbato, M., Precigout, G., Viornery, A., Bucher-Laurent, S., Benkhelifa-Ziyyat, S., Merten, O.W., Perea, J., and Galy, A. (2011). Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. Gene Ther *18*, 479-487

48. Charrier, S., Stockholm, D., Seye, K., Opolon, P., Taveau, M., Gross, D.A., Bucher-Laurent, S., Delenda, C., Vainchenker, W., Danos, O*., et al.* (2005). A lentiviral vector encoding the human Wiskott-Aldrich syndrome protein corrects immune and cytoskeletal defects in WASP knockout mice. Gene Ther *12*, 597-606

49. Chevalier-Mariette, C., Henry, I., Montfort, L., Capgras, S., Forlani, S., Muschler, J., and Nicolas, J.F. (2003). CpG content affects gene silencing in mice: evidence from novel transgenes. Genome Biol 4, R53

50. Coffin, J.M. (1992). Retroviral DNA integration. Dev Biol Stand *76*, 141-151

51. Coil, D.A., and Miller, A.D. (2004). Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. J Virol *78*, 10920-10926

52. Connor, J., Pak, C.C., and Schroit, A.J. (1994). Exposure of phosphatidylserine in the outer leaflet of human red blood cells. Relationship to cell density, cell age, and clearance by mononuclear cells. J Biol Chem *269*, 2399-2404

53. Cornils, K., Lange, C., Schambach, A., Brugman, M.H., Nowak, R., Lioznov, M., Baum, C., and Fehse, B. (2009). Stem cell marking with promotor-deprived self-inactivating retroviral vectors does not lead to induced clonal imbalance. Mol Ther *17*, 131-143

54. Costa, C., Albuisson, J., Le, T.H., Max-Audit, I., Dinh, K.T., Tosi, M., Goossens, M., and Pissard, S. (2005). Severe hemolytic anemia in a Vietnamese family, associated with novel mutations in the gene encoding for pyruvate kinase. Haematologica *90*, 25-30

55. Davidson, B.L., and McCray, P.B., Jr. (2011). Current prospects for RNA interference-based therapies. Nat Rev Genet 12, 329-340

56. Decaux, J.F., Antoine, B., and Kahn, A. (1989). Regulation of the expression of the L-type pyruvate kinase gene in adult rat hepatocytes in primary culture. J Biol Chem *264*, 11584-11590

57. de Jong, K., Emerson, R.K., Butler, J., Bastacky, J., Mohandas, N., and Kuypers, F.A. (2001). Short survival of phosphatidylserine-exposing red blood cells in murine sickle cell anemia. Blood 98, 1577-1584

58. Delenda, C. (2004). Lentiviral vectors: optimization of packaging, transduction and gene expression. J Gene Med *6 Suppl 1*, S125-138

59. Diez, A., Gilsanz, F., Martinez, J., Perez-Benavente, S., Meza, N.W., and Bautista, J.M. (2005). Life-threatening nonspherocytic hemolytic anemia in a patient with a null mutation in the PKLR gene and no compensatory PKM gene expression. Blood *106*, 1851-1856

60. Dolan, L.M., Ryan, M., and Moohan, J. (2002). Pyruvate kinase deficiency in pregnancy complicated by iron overload. BJOG 109, 844-846

61. Duan, J., and Antezana, M.A. (2003). Mammalian mutation pressure, synonymous codon choice, and mRNA degradation. J Mol Evol 57, 694-701

62. Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. J Virol *72*, 8463-8471

63. Dupre, L., Marangoni, F., Scaramuzza, S., Trifari, S., Hernandez, R.J., Aiuti, A., Naldini, L., and Roncarolo, M.G. (2006). Efficacy of gene therapy for Wiskott-Aldrich syndrome using a WAS promoter/cDNA-containing lentiviral vector and nonlethal irradiation. Hum Gene Ther *17*, 303-313

64. Ellis, J. (2005). Silencing and variegation of gammaretrovirus and lentivirus vectors. Hum Gene Ther *16*, 1241-1246

65. Emery, D.W. (2011). The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. Hum Gene Ther 22, 761-774

66. Escors, D., and Breckpot, K. (2010). Lentiviral vectors in gene therapy: their current status and future potential. Arch Immunol Ther Exp (Warsz) *58*, 107-119

67. Eshghi, S., Vogelezang, M.G., Hynes, R.O., Griffith, L.G., and Lodish, H.F. (2007). Alpha4beta1 integrin and erythropoietin mediate temporally distinct steps in erythropoiesis: integrins in red cell development. J Cell Biol *177*, 871-880

68. Felli, N., Pedini, F., Romania, P., Biffoni, M., Morsilli, O., Castelli, G., Santoro, S., Chicarella, S., Sorrentino, A., Peschle, C., et al. (2009). MicroRNA 223-dependent expression of LMO2 regulates normal erythropoiesis. Haematologica 94, 479-486

69. Fermo, E., Bianchi, P., Chiarelli, L.R., Cotton, F., Vercellati, C., Writzl, K., Baker, K., Hann, I., Rodwell, R., Valentini, G.*, et al.* (2005). Red cell pyruvate kinase deficiency: 17 new mutations of the PK-LR gene. Br J Haematol *129*, 839-846

70. Ferreira, P., Morais, L., Costa, R., Resende, C., Dias, C.P., Araujo, F., Costa, E., Barbot, J., and Vilarinho, A. (2000). Hydrops fetalis associated with erythrocyte pyruvate kinase deficiency. Eur J Pediatr *159*, 481-482

71. Fischer, A., Hacein-Bey-Abina, S., and Cavazzana-Calvo, M. (2010). 20 years of gene therapy for SCID. Nat Immunol *11*, 457-460

72. Fitton, L.A., and Bulfield, G. (1989). The liver/erythrocyte pyruvate kinase gene complex [Pk-1] in the mouse: structural gene mutations. Genet Res *53*, 105-110

73. Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M., and Naldini, L. (2000). Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. Nat Genet *25*, 217-222

74. Fontanellas, A., Mazurier, F., Landry, M., Taine, L., Morel, C., Larou, M., Daniel, J.Y., Montagutelli, X., de Salamanca, R.E., and de Verneuil, H. (2000). Reversion of hepatobiliary alterations By bone marrow transplantation in a murine model of erythropoietic protoporphyria. Hepatology *32*, 73-81

75. Fontanellas, A., Mendez, M., Mazurier, F., Cario-Andre, M., Navarro, S., Ged, C., Taine, L., Geronimi, F., Richard, E., Moreau-Gaudry, F.*, et al.* (2001). Successful therapeutic effect in a mouse model of erythropoietic protoporphyria by partial genetic correction and fluorescence-based selection of hematopoietic cells. Gene Ther *8*, 618-626

76. Forrester, W.C., Thompson, C., Elder, J.T., and Groudine, M. (1986). A developmentally stable chromatin structure in the human beta-globin gene cluster. Proc Natl Acad Sci U S A *83*, 1359-1363

77. Fortin, A., Cardon, L.R., Tam, M., Skamene, E., Stevenson, M.M., and Gros, P. (2001). Identification of a new malaria susceptibility locus (Char4) in recombinant congenic strains of mice. Proc Natl Acad Sci U S A *98*, 10793-10798

78. Fothergill-Gilmore, L.A., and Michels, P.A. (1993). Evolution of glycolysis. Prog Biophys Mol Biol *59*, 105-235

79. Fraser, P., and Grosveld, F. (1998). Locus control regions, chromatin activation and transcription. Curr Opin Cell Biol *10*, 361-365

80. Frecha, C., Fusil, F., Cosset, F.L., and Verhoeyen, E. (2011). In vivo gene delivery into hCD34+ cells in a humanized mouse model. Methods Mol Biol *737*, 367-390

81. Frecha, C., Toscano, M.G., Costa, C., Saez-Lara, M.J., Cosset, F.L., Verhoeyen, E., and Martin, F. (2008). Improved lentiviral vectors for Wiskott-Aldrich syndrome gene therapy mimic endogenous expression profiles throughout haematopoiesis. Gene Ther *15*, 930-941

82. Fung, R.H., Keung, Y.K., and Chung, G.S. (1969). Screening of pyruvate kinase deficiency and G6PD deficiency in Chinese newborn in Hong Kong. Arch Dis Child *44*, 373-376

83. Galy, A., Roncarolo, M.G., and Thrasher, A.J. (2008). Development of lentiviral gene therapy for Wiskott Aldrich syndrome. Expert Opin Biol Ther *8*, 181-190

84. García-Gómez, M., Quintana-Bustamante, O., García-Bravo, M., Navarro, S., Gárate, Z. and Segovia, J.C. (2013). Gene therapy for erythroid metabolic inherited diseases. "Gene therapy-Tools and potential applications. Book edited by Francisco Molina, ISBN 979-953-51-1014-9

85. Gaspar, H.B., Bjorkegren, E., Parsley, K., Gilmour, K.C., King, D., Sinclair, J., Zhang, F., Giannakopoulos, A., Adams, S., Fairbanks, L.D*., et al.* (2006). Successful reconstitution of immunity in ADA-SCID by stem cell gene therapy following cessation of PEG-ADA and use of mild preconditioning. Mol Ther *14*, 505-513

86. Gaspar, H.B., Parsley, K.L., Howe, S., King, D., Gilmour, K.C., Sinclair, J., Brouns, G., Schmidt, M., Von Kalle, C., Barington, T*., et al.* (2004). Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. Lancet *364*, 2181-2187

87. Gerolami, R., Uch, R., Jordier, F., Chapel, S., Bagnis, C., Brechot, C., and Mannoni, P. (2000). Gene transfer to hepatocellular carcinoma: transduction efficacy and transgene expression kinetics by using retroviral and lentiviral vectors. Cancer Gene Ther 7, 1286-1292

88. Giger, U., Mason, G.D., and Wang, P. (1991). Inherited erythrocyte pyruvate kinase deficiency in a beagle dog. Vet Clin Pathol *20*, 83-86

89. Giger, U., and Noble, N.A. (1991). Determination of erythrocyte pyruvate kinase deficiency in Basenjis with chronic hemolytic anemia. J Am Vet Med Assoc *198*, 1755-1761

90. Gilsanz, F., Vega, M.A., Gomez-Castillo, E., Ruiz-Balda, J.A., and Omenaca, F. (1993). Fetal anaemia due to pyruvate kinase deficiency. Arch Dis Child *69*, 523-524

91. Ginn, S.L., Fleming, J., Rowe, P.B., and Alexander, I.E. (2003). Promoter interference mediated by the U3 region in early-generation HIV-1-derived lentivirus vectors can influence detection of transgene expression in a cell-type and species-specific manner. Hum Gene Ther *14*, 1127-1137

92. Gonzalez-Murillo, A., Lozano, M.L., Alvarez, L., Jacome, A., Almarza, E., Navarro, S., Segovia, J.C., Hanenberg, H., Guenechea, G., Bueren, J.A., et al. (2010). Development of lentiviral

vectors with optimized transcriptional activity for the gene therapy of patients with Fanconi anemia. Hum Gene Ther 21, 623-630

93. Gonzalez-Murillo, A., Lozano, M.L., Montini, E., Bueren, J.A., and Guenechea, G. (2008). Unaltered repopulation properties of mouse hematopoietic stem cells transduced with lentiviral vectors. Blood *112*, 3138-3147

94. Grez, M., Reichenbach, J., Schwable, J., Seger, R., Dinauer, M.C., and Thrasher, A.J. (2011). Gene therapy of chronic granulomatous disease: the engraftment dilemma. Mol Ther *19*, 28-35

95. Grosveld, F. (1999). Activation by locus control regions? Curr Opin Genet Dev *9*, 152-157

96. Guenechea, G., Gan, O.I., Inamitsu, T., Dorrell, C., Pereira, D.S., Kelly, M., Naldini, L., and Dick, J.E. (2000). Transduction of human CD34+ CD38- bone marrow and cord blood-derived SCID-repopulating cells with third-generation lentiviral vectors. Mol Ther *1*, 566-573

97. Guguen-Guillouzo, C., Szajnert, M.F., Marie, J., Delain, D., and Schapira, F. (1977). Differentiation in vivo and in vitro of pyruvate kinase isozymes in rat muscle. Biochimie *59*, 65-71

98. Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K*., et al.* (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest *118*, 3132-3142

99. Hacein-Bey-Abina, S., Hauer, J., Lim, A., Picard, C., Wang, G.P., Berry, C.C., Martinache, C., Rieux-Laucat, F., Latour, S., Belohradsky, B.H*., et al.* (2010). Efficacy of gene therapy for X-linked severe combined immunodeficiency. N Engl J Med *363*, 355-364

100. Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 302, 415-419

101. Hamanaka, R.B., and Chandel, N.S. (2011). Cell biology. Warburg effect and redox balance. Science 334, 1219-1220

102. Hanawa, H., Hargrove, P.W., Kepes, S., Srivastava, D.K., Nienhuis, A.W., and Persons, D.A. (2004). Extended beta-globin locus control region elements promote consistent therapeutic expression of a gamma-globin lentiviral vector in murine beta-thalassemia. Blood *104*, 2281-2290

103. Hanawa, H., Kelly, P.F., Nathwani, A.C., Persons, D.A., Vandergriff, J.A., Hargrove, P., Vanin, E.F., and Nienhuis, A.W. (2002). Comparison of various envelope proteins for their ability to pseudotype lentiviral vectors and transduce primitive hematopoietic cells from human blood. Mol Ther *5*, 242-251

104. Hargrove, P.W., Kepes, S., Hanawa, H., Obenauer, J.C., Pei, D., Cheng, C., Gray, J.T., Neale, G., and Persons, D.A. (2008). Globin lentiviral vector insertions can perturb the expression of endogenous genes in beta-thalassemic hematopoietic cells. Mol Ther *16*, 525-533

105. Hattangadi, S.M., Wong, P., Zhang, L., Flygare, J., and Lodish, H.F. (2011). From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. Blood *118*, 6258-6268

106. Hematti, P., Hong, B.K., Ferguson, C., Adler, R., Hanawa, H., Sellers, S., Holt, I.E., Eckfeldt, C.E., Sharma, Y., Schmidt, M*., et al.* (2004). Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS Biol *2*, e423

107. Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D*., et al.* (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest *118*, 3143-3150

108. Huston, M.W., van Til, N.P., Visser, T.P., Arshad, S., Brugman, M.H., Cattoglio, C., Nowrouzi, A., Li, Y., Schambach, A., Schmidt, M., et al. (2011). Correction of murine SCID-X1 by lentiviral gene therapy using a codon-optimized IL2RG gene and minimal pretransplant conditioning. Mol Ther 19, 1867-1877.

109. Iwakuma, T., Cui, Y., and Chang, L.J. (1999). Self-inactivating lentiviral vectors with U3 and U5 modifications. Virology *261*, 120-132

110. Jacobasch, G., and Rapoport, S.M. (1996). Hemolytic anemias due to erythrocyte enzyme deficiencies. Mol Aspects Med 17, 143-170

111. Jacome, A., Navarro, S., Rio, P., Yanez, R.M., Gonzalez-Murillo, A., Lozano, M.L., Lamana, M.L., Sevilla, J., Olive, T., Diaz-Heredia, C*., et al.* (2009). Lentiviral-mediated genetic correction of hematopoietic and mesenchymal progenitor cells from Fanconi anemia patients. Mol Ther *17*, 1083-1092

112. Jin, L., Zeng, H., Chien, S., Otto, K.G., Richard, R.E., Emery, D.W., and Blau, C.A. (2000). In vivo selection using a cell-growth switch. Nat Genet *26*, 64-66

113. Johnnidis, J.B., Harris, M.H., Wheeler, R.T., Stehling-Sun, S., Lam, M.H., Kirak, O., Brummelkamp, T.R., Fleming, M.D., and Camargo, F.D. (2008). Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. Nature 451, 1125-1129

114. Johnston, J.M., Denning, G., Doering, C.B., and Spencer, H.T. (2012). Generation of an optimized lentiviral vector encoding a high-expression factor VIII transgene for gene therapy of hemophilia A. Gene Ther [Epub ahead of print]

115. Kahn, A., and Marie, J. (1982). Pyruvate kinases from human erythrocytes and liver. Methods Enzymol *90 Pt E*, 131-140

116. Kanno, H., Fujii, H., Hirono, A., and Miwa, S. (1991). cDNA cloning of human R-type pyruvate kinase and identification of a single amino acid substitution (Thr384----Met) affecting enzymatic stability in a pyruvate kinase variant (PK Tokyo) associated with hereditary hemolytic anemia. Proc Natl Acad Sci U S A *88*, 8218-8221

117. Kanno, H., Fujii, H., and Miwa, S. (1992). Structural analysis of human pyruvate kinase L-gene and identification of the promoter activity in erythroid cells. Biochem Biophys Res Commun *188*, 516-523

118. Kanno, H., Morimoto, M., Fujii, H., Tsujimura, T., Asai, H., Noguchi, T., Kitamura, Y., and Miwa, S. (1995). Primary structure of murine red blood cell-type pyruvate kinase (PK) and molecular characterization of PK deficiency identified in the CBA strain. Blood *86*, 3205-3210

119. Kanno, H., Utsugisawa, T., Aizawa, S., Koizumi, T., Aisaki, K., Hamada, T., Ogura, H., and Fujii, H. (2007). Transgenic rescue of hemolytic anemia due to red blood cell pyruvate kinase deficiency. Haematologica *92*, 731-737

120. Kanno, H., Wei, D.C., Chan, L.C., Mizoguchi, H., Ando, M., Nakahata, T., Narisawa, K., Fujii, H., and Miwa, S. (1994). Hereditary hemolytic anemia caused by diverse point mutations of pyruvate kinase gene found in Japan and Hong Kong. Blood *84*, 3505-3509

121. Kay, M.M. (1992). Molecular mapping of human band 3 aging antigenic sites and active amino acids using synthetic peptides. J Protein Chem *11*, 595-602

122. Keitt, A.S. (1966). Pyruvate kinase deficiency and related disorders of red cell glycolysis. Am J Med *41*, 762-785

123. Kerenyi, M.A., Grebien, F., Gehart, H., Schifrer, M., Artaker, M., Kovacic, B., Beug, H., Moriggl, R., and Mullner, E.W. (2008). Stat5 regulates cellular iron uptake of erythroid cells via IRP-2 and TfR-1. Blood 112, 3878-3888

124. Kerenyi, M.A., and Orkin, S.H. (2010). Networking erythropoiesis. J Exp Med *207*, 2537-2541

125. Kingsman, S.M., Mitrophanous, K., and Olsen, J.C. (2005). Potential oncogene activity of the woodchuck hepatitis post-transcriptional regulatory element (WPRE). Gene Ther *12*, 3-4

126. Kozak, M. (1996). Interpreting cDNA sequences: some insights from studies on translation. Mamm Genome *7*, 563-574

127. Lakomek, M., Schroter, W., De Maeyer, G., and Winkler, H. (1989). On the diagnosis of erythrocyte enzyme defects in the presence of high reticulocyte counts. Br J Haematol *72*, 445-451

128. Lakomek, M., Winkler, H., Pekrun, A., Kruger, N., Sander, M., Huppke, P., and Schroter, W. (1994). Erythrocyte pyruvate kinase deficiency. The influence of physiologically important metabolites on the function of normal and defective enzymes. Enzyme Protein 48, 149-163

129. Lenzner, C., Nurnberg, P., Jacobasch, G., Gerth, C., and Thiele, B.J. (1997). Molecular analysis of 29 pyruvate kinase-deficient patients from central Europe with hereditary hemolytic anemia. Blood *89*, 1793-1799

130. Lenzner, C., Nurnberg, P., Thiele, B.J., Reis, A., Brabec, V., Sakalova, A., and Jacobasch, G. (1994). Mutations in the pyruvate kinase L gene in patients with hereditary hemolytic anemia. Blood *83*, 2817-2822

131. Lewis, P.F., and Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. J Virol *68*, 510-516

132. Li, Q., Peterson, K.R., Fang, X., and Stamatoyannopoulos, G. (2002). Locus control regions. Blood *100*, 3077-3086

133. Liang, S., Moghimi, B., Yang, T.P., Strouboulis, J., and Bungert, J. (2008). Locus control region mediated regulation of adult beta-globin gene expression. J Cell Biochem *105*, 9-16

134. Lisowski, L., and Sadelain, M. (2008). Current status of globin gene therapy for the treatment of beta-thalassaemia. Br J Haematol *141*, 335-345

135. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408

136. Lozzio, C.B., and Lozzio, B.B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. Blood 45, 321-334

137. Lunt, S.Y., and Vander Heiden, M.G. (2011). Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. Annu Rev Cell Dev Biol *27*, 441-464

138. Machado, P., Manco, L., Gomes, C., Mendes, C., Fernandes, N., Salome, G., Sitoe, L., Chibute, S., Langa, J., Ribeiro, L., et al. (2012). Pyruvate kinase deficiency in sub-Saharan Africa:

identification of a highly frequent missense mutation (G829A;Glu277Lys) and association with malaria. PLoS One 7, e47071

139. Machado, P., Pereira, R., Rocha, A.M., Manco, L., Fernandes, N., Miranda, J., Ribeiro, L., do Rosario, V.E., Amorim, A., Gusmao, L*., et al.* (2010). Malaria: looking for selection signatures in the human PKLR gene region. Br J Haematol *149*, 775-784

140. Malik, P., Arumugam, P.I., Yee, J.K., and Puthenveetil, G. (2005). Successful correction of the human Cooley's anemia beta-thalassemia major phenotype using a lentiviral vector flanked by the chicken hypersensitive site 4 chromatin insulator. Ann N Y Acad Sci 1054, 238-249

141. Manco, L., Ribeiro, M.L., Maximo, V., Almeida, H., Costa, A., Freitas, O., Barbot, J., Abade, A., and Tamagnini, G. (2000). A new PKLR gene mutation in the R-type promoter region affects the gene transcription causing pyruvate kinase deficiency. Br J Haematol *110*, 993-997

142. Mantovani, J., Charrier, S., Eckenberg, R., Saurin, W., Danos, O., Perea, J., and Galy, A. (2009). Diverse genomic integration of a lentiviral vector developed for the treatment of Wiskott-Aldrich syndrome. J Gene Med *11*, 645-654

143. Marinkovic, D., Zhang, X., Yalcin, S., Luciano, J.P., Brugnara, C., Huber, T., and Ghaffari, S. (2007). Foxo3 is required for the regulation of oxidative stress in erythropoiesis. J Clin Invest *117*, 2133-2144

144. Marshall, S.R., Saunders, P.W., Hamilton, P.J., and Taylor, P.R. (2003). The dangers of iron overload in pyruvate kinase deficiency. Br J Haematol *120*, 1090-1091

145. Martin, P., and Papayannopoulou, T. (1982). HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression. Science 216, 1233-1235

146. Matrai, J., Chuah, M.K., and VandenDriessche, T. (2010). Recent advances in lentiviral vector development and applications. Mol Ther 18, 477-490

147. Max-Audit, I., Eleouet, J.F., and Romeo, P.H. (1993). Transcriptional regulation of the pyruvate kinase erythroid-specific promoter. J Biol Chem 268, 5431-5437

148. Max-Audit, I., Kechemir, D., Mitjavila, M.T., Vainchenker, W., Rotten, D., and Rosa, R. (1988). Pyruvate kinase synthesis and degradation by normal and pathologic cells during erythroid maturation. Blood *72*, 1039-1044

149. Max-Audit, I., Testa, U., Kechemir, D., Titeux, M., Vainchenker, W., and Rosa, R. (1984). Pattern of pyruvate kinase isozymes in erythroleukemia cell lines and in normal human erythroblasts. Blood *64*, 930-936

150. May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K.M., Luzzatto, L., and Sadelain, M. (2000). Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. Nature *406*, 82-86

151. May, C., Rivella, S., Chadburn, A., and Sadelain, M. (2002). Successful treatment of murine beta-thalassemia intermedia by transfer of the human beta-globin gene. Blood *99*, 1902-1908

152. Meza, N.W., Alonso-Ferrero, M.E., Navarro, S., Quintana-Bustamante, O., Valeri, A., Garcia-Gomez, M., Bueren, J.A., Bautista, J.M., and Segovia, J.C. (2009). Rescue of pyruvate kinase deficiency in mice by gene therapy using the human isoenzyme. Mol Ther *17*, 2000-2009

153. Meza, N.W., Quintana-Bustamante, O., Puyet, A., Rio, P., Navarro, S., Diez, A., Bueren, J.A., Bautista, J.M., and Segovia, J.C. (2007). In vitro and in vivo expression of human erythrocyte pyruvate kinase in erythroid cells: a gene therapy approach. Hum Gene Ther *18*, 502-514

154. Miccio, A., Cesari, R., Lotti, F., Rossi, C., Sanvito, F., Ponzoni, M., Routledge, S.J., Chow, C.M., Antoniou, M.N., and Ferrari, G. (2008). In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of beta-thalassemia. Proc Natl Acad Sci U S A *105*, 10547-10552

155. Min-Oo, G., Fortin, A., Tam, M.F., Gros, P., and Stevenson, M.M. (2004). Phenotypic expression of pyruvate kinase deficiency and protection against malaria in a mouse model. Genes Immun *5*, 168-175

156. Min-Oo, G., Fortin, A., Tam, M.F., Nantel, A., Stevenson, M.M., and Gros, P. (2003). Pyruvate kinase deficiency in mice protects against malaria. Nat Genet *35*, 357-362

157. Min-Oo, G., Tam, M., Stevenson, M.M., and Gros, P. (2007). Pyruvate kinase deficiency: correlation between enzyme activity, extent of hemolytic anemia and protection against malaria in independent mouse mutants. Blood Cells Mol Dis *39*, 63-69

158. Mitchell, R.S., Beitzel, B.F., Schroder, A.R., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol *2*, E234

159. Miwa, S., and Fujii, H. (1985). Molecular aspects of erythroenzymopathies associated with hereditary hemolytic anemia. Am J Hematol *19*, 293-305

160. Miwa, S., Nakashima, K., Ariyoshi, K., Shinohara, K., and Oda, E. (1975). Four new pyruvate kinase (PK) variants and a classical PK deficiency. Br J Haematol *29*, 157-169

161. Miyoshi, H., Blomer, U., Takahashi, M., Gage, F.H., and Verma, I.M. (1998). Development of a self-inactivating lentivirus vector. J Virol *72*, 8150-8157

162. Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M.H., Schambach, A., Charrier, S., Galy, A., Thrasher, A.J.*, et al.* (2009). Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. Mol Ther *17*, 1919-1928

163. Montiel-Equihua, C.A., Zhang, L., Knight, S., Saadeh, H., Scholz, S., Carmo, M., Alonso-Ferrero, M.E., Blundell, M.P., Monkeviciute, A., Schulz, R.*, et al.* (2012). The beta-globin locus control region in combination with the EF1alpha short promoter allows enhanced lentiviral vector-mediated erythroid gene expression with conserved multilineage activity. Mol Ther *20*, 1400-1409

164. Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., Sergi Sergi, L., Benedicenti, F., Ambrosi, A., Di Serio, C.*, et al.* (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat Biotechnol *24*, 687-696

165. Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Bartholomae, C.C., Ranzani, M., Benedicenti, F., Sergi, L.S., Ambrosi, A., Ponzoni, M., et al. (2009). The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. J Clin Invest 119, 964-975

166. Moreau-Gaudry, F., Xia, P., Jiang, G., Perelman, N.P., Bauer, G., Ellis, J., Surinya, K.H., Mavilio, F., Shen, C.K., and Malik, P. (2001). High-level erythroid-specific gene expression in primary

human and murine hematopoietic cells with self-inactivating lentiviral vectors. Blood *98*, 2664-2672

167. Morimoto, M., Kanno, H., Asai, H., Tsujimura, T., Fujii, H., Moriyama, Y., Kasugai, T., Hirono, A., Ohba, Y., Miwa, S.*, et al.* (1995). Pyruvate kinase deficiency of mice associated with nonspherocytic hemolytic anemia and cure of the anemia by marrow transplantation without host irradiation. Blood *86*, 4323-4330

168. Morris, J.C., Conerly, M., Thomasson, B., Storek, J., Riddell, S.R., and Kiem, H.P. (2004). Induction of cytotoxic T-lymphocyte responses to enhanced green and yellow fluorescent proteins after myeloablative conditioning. Blood 103, 492-499

169. Mostoslavsky, G., Kotton, D.N., Fabian, A.J., Gray, J.T., Lee, J.S., and Mulligan, R.C. (2005). Efficiency of transduction of highly purified murine hematopoietic stem cells by lentiviral and oncoretroviral vectors under conditions of minimal in vitro manipulation. Mol Ther *11*, 932-940

170. Mullerova, J. (2004). Use of Recombinant Congenic Strains in Mapping Disease-Modifying Genes. News in Physiological Sciences 19, 105-109

171. Nakashima, K., Miwa, S., Fujii, H., Shinohara, K., Yamauchi, K., Tsuji, Y., and Yanai, M. (1977). Characterization of pyruvate kinase from the liver of a patient with aberrant erythrocyte pyruvate kinase, PK Nagasaki. J Lab Clin Med *90*, 1012-1020

172. Naldini, L. (2011). Ex vivo gene transfer and correction for cell-based therapies. Nat Rev Genet *12*, 301-315

173. Naldini, L., Blomer, U., Gage, F.H., Trono, D., and Verma, I.M. (1996). Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc Natl Acad Sci U S A *93*, 11382-11388

174. Navarro, S., Meza, N.W., Quintana-Bustamante, O., Casado, J.A., Jacome, A., McAllister, K., Puerto, S., Surralles, J., Segovia, J.C., and Bueren, J.A. (2006). Hematopoietic dysfunction in a mouse model for Fanconi anemia group D1. Mol Ther 14, 525-535

175. Navas, P.A., Peterson, K.R., Li, Q., McArthur, M., and Stamatoyannopoulos, G. (2001). The 5'HS4 core element of the human beta-globin locus control region is required for high-level globin gene expression in definitive but not in primitive erythropoiesis. J Mol Biol *312*, 17-26

176. Neubauer, B., Lakomek, M., Winkler, H., Parke, M., Hofferbert, S., and Schroter, W. (1991). Point mutations in the L-type pyruvate kinase gene of two children with hemolytic anemia caused by pyruvate kinase deficiency. Blood *77*, 1871-1875

177. Ney, P.A. (2011). Normal and disordered reticulocyte maturation. Curr Opin Hematol 18, 152-157

178. Nijhof, W., Wierenga, P.K., Staal, G.E., and Jansen, G. (1984). Changes in activities and isozyme patterns of glycolytic enzymes during erythroid differentiation in vitro. Blood *64*, 607-613

179. Noguchi, T., Yamada, K., Inoue, H., Matsuda, T., and Tanaka, T. (1987). The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters. J Biol Chem *262*, 14366-14371

180. Oertel, M., Rosencrantz, R., Chen, Y.Q., Thota, P.N., Sandhu, J.S., Dabeva, M.D., Pacchia, A.L., Adelson, M.E., Dougherty, J.P., and Shafritz, D.A. (2003). Repopulation of rat liver by fetal

hepatoblasts and adult hepatocytes transduced ex vivo with lentiviral vectors. Hepatology 37, 994-1005

181. Oski, F.A., Marshall, B.E., Cohen, P.J., Sugerman, H.J., and Miller, L.D. (1971). The role of the left-shifted or right-shifted oxygen-hemoglobin equilibrium curve. Ann Intern Med *74*, 44-46

182. Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kuhlcke, K., Schilz, A., Kunkel, H.*, et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med *12*, 401-409

183. Palstra, R.J., de Laat, W., and Grosveld, F. (2008). Beta-globin regulation and long-range interactions. Adv Genet *61*, 107-142

184. Passow, H. (1986). Molecular aspects of band 3 protein-mediated anion transport across the red blood cell membrane. Rev Physiol Biochem Pharmacol *103*, 61-203

185. Paulson, R.F., Shi, L., and Wu, D.C. (2011). Stress erythropoiesis: new signals and new stress progenitor cells. Curr Opin Hematol *18*, 139-145

186. Pawliuk, R., Westerman, K.A., Fabry, M.E., Payen, E., Tighe, R., Bouhassira, E.E., Acharya, S.A., Ellis, J., London, I.M., Eaves, C.J.*, et al.* (2001). Correction of sickle cell disease in transgenic mouse models by gene therapy. Science *294*, 2368-2371

187. Perry, J.M., Harandi, O.F., Porayette, P., Hegde, S., Kannan, A.K., and Paulson, R.F. (2009). Maintenance of the BMP4-dependent stress erythropoiesis pathway in the murine spleen requires hedgehog signaling. Blood *113*, 911-918

188. Persons, D.A., Hargrove, P.W., Allay, E.R., Hanawa, H., and Nienhuis, A.W. (2003). The degree of phenotypic correction of murine beta -thalassemia intermedia following lentiviral-mediated transfer of a human gamma-globin gene is influenced by chromosomal position effects and vector copy number. Blood *101*, 2175-2183

189. Perumbeti, A., Higashimoto, T., Urbinati, F., Franco, R., Meiselman, H.J., Witte, D., and Malik, P. (2009). A novel human gamma-globin gene vector for genetic correction of sickle cell anemia in a humanized sickle mouse model: critical determinants for successful correction. Blood *114*, 1174-1185

190. Pestina, T.I., Hargrove, P.W., Jay, D., Gray, J.T., Boyd, K.M., and Persons, D.A. (2009). Correction of murine sickle cell disease using gamma-globin lentiviral vectors to mediate high-level expression of fetal hemoglobin. Mol Ther *17*, 245-252

191. Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res *29*, e45

192. Pfeifer, A., Ikawa, M., Dayn, Y., and Verma, I.M. (2002). Transgenesis by lentiviral vectors: lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. Proc Natl Acad Sci U S A *99*, 2140-2145

193. Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001). Cooperation and competition in the evolution of ATP-producing pathways. Science *292*, 504-507

194. Puthenveetil, G., Scholes, J., Carbonell, D., Qureshi, N., Xia, P., Zeng, L., Li, S., Yu, Y., Hiti, A.L., Yee, J.K.*, et al.* (2004). Successful correction of the human beta-thalassemia major phenotype using a lentiviral vector. Blood *104*, 3445-3453

195. Raich, N., and Romeo, P.H. (1993). Erythroid regulatory elements. Stem Cells *11*, 95-104

196. Ramezani, A., and Hawley, R.G. (2010). Strategies to insulate lentiviral vector-expressed transgenes. Methods Mol Biol *614*, 77-100

197. Ramezani, A., Hawley, T.S., and Hawley, R.G. (2003). Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. Blood *101*, 4717-4724

198. Richard, E., Geronimi, F., Lalanne, M., Ged, C., Redonnet-Vernhet, I., Lamrissi-Garcia, I., Gerson, S.L., de Verneuil, H., and Moreau-Gaudry, F. (2003). A bicistronic SIN-lentiviral vector containing G156A MGMT allows selection and metabolic correction of hematopoietic protoporphyric cell lines. J Gene Med *5*, 737-747

199. Richard, E., Mendez, M., Mazurier, F., Morel, C., Costet, P., Xia, P., Fontanellas, A., Geronimi, F., Cario-Andre, M., Taine, L.*, et al.* (2001). Gene therapy of a mouse model of protoporphyria with a self-inactivating erythroid-specific lentiviral vector without preselection. Mol Ther *4*, 331-338

200. Richard, R.E., Weinreich, M., Chang, K.H., Ieremia, J., Stevenson, M.M., and Blau, C.A. (2004). Modulating erythrocyte chimerism in a mouse model of pyruvate kinase deficiency. Blood *103*, 4432-4439

201. Richard, R.E., Wood, B., Zeng, H., Jin, L., Papayannopoulou, T., and Blau, C.A. (2000). Expansion of genetically modified primary human hemopoietic cells using chemical inducers of dimerization. Blood *95*, 430-436

202. Richmond, T.D., Chohan, M., and Barber, D.L. (2005). Turning cells red: signal transduction mediated by erythropoietin. Trends Cell Biol *15*, 146-155

203. Rigden, D.J., Phillips, S.E., Michels, P.A., and Fothergill-Gilmore, L.A. (1999). The structure of pyruvate kinase from Leishmania mexicana reveals details of the allosteric transition and unusual effector specificity. J Mol Biol *291*, 615-635

204. Rio, P., Meza, N.W., Gonzalez-Murillo, A., Navarro, S., Alvarez, L., Surralles, J., Castella, M., Guenechea, G., Segovia, J.C., Hanenberg, H.*, et al.* (2008). In vivo proliferation advantage of genetically corrected hematopoietic stem cells in a mouse model of Fanconi anemia FA-D1. Blood *112*, 4853-4861

205. Rivella, S. (2012). The role of ineffective erythropoiesis in non-transfusion-dependent thalassemia. Blood Rev *26 Suppl 1*, S12-15

206. Rivella, S., May, C., Chadburn, A., Riviere, I., and Sadelain, M. (2003). A novel murine model of Cooley anemia and its rescue by lentiviral-mediated human beta-globin gene transfer. Blood *101*, 2932-2939

207. Robert-Richard, E., Moreau-Gaudry, F., Lalanne, M., Lamrissi-Garcia, I., Cario-Andre, M., Guyonnet-Duperat, V., Taine, L., Ged, C., and de Verneuil, H. (2008). Effective gene therapy of mice with congenital erythropoietic porphyria is facilitated by a survival advantage of corrected erythroid cells. Am J Hum Genet *82*, 113-124

208. Roncarolo, M.G., and Battaglia, M. (2007). Regulatory T-cell immunotherapy for tolerance to self antigens and alloantigens in humans. Nat Rev Immunol *7*, 585-598

209. Roselli, E.A., Mezzadra, R., Frittoli, M.C., Maruggi, G., Biral, E., Mavilio, F., Mastropietro, F., Amato, A., Tonon, G., Refaldi, C., et al. (2010). Correction of beta-thalassemia major by gene transfer in haematopoietic progenitors of pediatric patients. EMBO Mol Med 2, 315-328

210. Rossi, D.J., Bryder, D., Seita, J., Nussenzweig, A., Hoeijmakers, J., and Weissman, I.L. (2007). Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. Nature *447*, 725-729

211. Sadelain, M., Chang, A., and Lisowski, L. (2009). Supplying clotting factors from hematopoietic stem cell-derived erythroid and megakaryocytic lineage cells. Mol Ther *17*, 1994-1999

212. Sakuma, T., Barry, M.A., and Ikeda, Y. (2012). Lentiviral vectors: basic to translational. Biochem J *443*, 603-618

213. Sastry, L., Johnson, T., Hobson, M.J., Smucker, B., and Cornetta, K. (2002). Titering lentiviral vectors: comparison of DNA, RNA and marker expression methods. Gene Ther *9*, 1155-1162

214. Satoh, H., Tani, K., Yoshida, M.C., Sasaki, M., Miwa, S., and Fujii, H. (1988). The human liver-type pyruvate kinase (PKL) gene is on chromosome 1 at band q21. Cytogenet Cell Genet *47*, 132-133

215. Schambach, A., Galla, M., Maetzig, T., Loew, R., and Baum, C. (2007). Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors. Mol Ther *15*, 1167-1173

216. Schambach, A., Swaney, W.P., and van der Loo, J.C. (2009). Design and production of retro- and lentiviral vectors for gene expression in hematopoietic cells. Methods Mol Biol *506*, 191-205

217. Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110*, 521-529

218. Sedano, I.B., Rothlisberger, B., Deleze, G., Ottiger, C., Panchard, M.A., Spahr, A., Hergersberg, M., Burgi, W., and Huber, A. (2004). PK Aarau: first homozygous nonsense mutation causing pyruvate kinase deficiency. Br J Haematol *127*, 364-366

219. Sinn, P.L., Sauter, S.L., and McCray, P.B., Jr. (2005). Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production. Gene Ther *12*, 1089-1098

220. Skelly, B.J., Wallace, M., Rajpurohit, Y.R., Wang, P., and Giger, U. (1999). Identification of a 6 base pair insertion in West Highland White Terriers with erythrocyte pyruvate kinase deficiency. Am J Vet Res *60*, 1169-1172

221. Stein, S., Ott, M.G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., Schmidt, M., Kramer, A., Schwable, J., Glimm, H*., et al.* (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat Med *16*, 198-204

222. Stripecke, R., Carmen Villacres, M., Skelton, D., Satake, N., Halene, S., and Kohn, D. (1999). Immune response to green fluorescent protein: implications for gene therapy. Gene Ther 6, 1305-1312

223. Takatu, A., Nash, R.A., Zaucha, J.M., Little, M.T., Georges, G.E., Sale, G.E., Zellmer, E., Kuhr, C.S., Lothrop, C.D., Jr., and Storb, R. (2003). Adoptive immunotherapy to increase the level of donor hematopoietic chimerism after nonmyeloablative marrow transplantation for severe canine hereditary hemolytic anemia. Biol Blood Marrow Transplant *9*, 674-682

224. Takegawa, S., Fujii, H., and Miwa, S. (1983). Change of pyruvate kinase isozymes from M2- to L-type during development of the red cell. Br J Haematol *54*, 467-474

225. Takenaka, M., Noguchi, T., Sadahiro, S., Hirai, H., Yamada, K., Matsuda, T., Imai, E., and Tanaka, T. (1991). Isolation and characterization of the human pyruvate kinase M gene. Eur J Biochem *198*, 101-106

226. Tan, M., Qing, K., Zhou, S., Yoder, M.C., and Srivastava, A. (2001). Adeno-associated virus 2-mediated transduction and erythroid lineage-restricted long-term expression of the human beta-globin gene in hematopoietic cells from homozygous beta-thalassemic mice. Mol Ther *3*, 940-946

227. Tani, K., Yoshikubo, T., Ikebuchi, K., Takahashi, K., Tsuchiya, T., Takahashi, S., Shimane, M., Ogura, H., Tojo, A., Ozawa, K*., et al.* (1994). Retrovirus-mediated gene transfer of human pyruvate kinase (PK) cDNA into murine hematopoietic cells: implications for gene therapy of human PK deficiency. Blood *83*, 2305-2310

228. Tanphaichitr, V.S., Suvatte, V., Issaragrisil, S., Mahasandana, C., Veerakul, G., Chongkolwatana, V., Waiyawuth, W., and Ideguchi, H. (2000). Successful bone marrow transplantation in a child with red blood cell pyruvate kinase deficiency. Bone Marrow Transplant *26*, 689-690

229. Throm, R.E., Ouma, A.A., Zhou, S., Chandrasekaran, A., Lockey, T., Greene, M., De Ravin, S.S., Moayeri, M., Malech, H.L., Sorrentino, B.P*., et al.* (2009). Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection. Blood *113*, 5104-5110

230. Toscano, M.G., Romero, Z., Munoz, P., Cobo, M., Benabdellah, K., and Martin, F. (2011). Physiological and tissue-specific vectors for treatment of inherited diseases. Gene Ther *18*, 117-127

231. Trobridge, G.D., Beard, B.C., Wu, R.A., Ironside, C., Malik, P., and Kiem, H.P. (2012). Stem cell selection in vivo using foamy vectors cures canine pyruvate kinase deficiency. PLoS One *7*, e45173

232. Trono, D. (2003). Virology. Picking the right spot. Science *300*, 1670-1671

233. Tsujimoto, Y. (1997). Apoptosis and necrosis: intracellular ATP level as a determinant for cell death modes. Cell Death Differ *4*, 429-434

234. Valentine, W.N., Tanaka, K.R., and Miwa, S. (1961). A specific erythrocyte glycolytic enzyme defect (pyruvate kinase) in three subjects with congenital non-spherocytic hemolytic anemia. Trans Assoc Am Physicians *74*, 100-110

235. Valentini, G., Chiarelli, L.R., Fortin, R., Dolzan, M., Galizzi, A., Abraham, D.J., Wang, C., Bianchi, P., Zanella, A., and Mattevi, A. (2002). Structure and function of human erythrocyte pyruvate kinase. Molecular basis of nonspherocytic hemolytic anemia. J Biol Chem *277*, 23807-23814

236. VandenDriessche, T., Thorrez, L., Naldini, L., Follenzi, A., Moons, L., Berneman, Z., Collen, D., and Chuah, M.K. (2002). Lentiviral vectors containing the human immunodeficiency virus type-1 central polypurine tract can efficiently transduce nondividing hepatocytes and antigen-presenting cells in vivo. Blood 100, 813-822

237. van Til, N.P., de Boer, H., Mashamba, N., Wabik, A., Huston, M., Visser, T.P., Fontana, E., Poliani, P.L., Cassani, B., Zhang, F., et al. (2012). Correction of murine Rag2 severe combined

immunodeficiency by lentiviral gene therapy using a codon-optimized RAG2 therapeutic transgene. Mol Ther 20, 1968-1980

238. van Wijk, R., Huizinga, E.G., van Wesel, A.C., van Oirschot, B.A., Hadders, M.A., and van Solinge, W.W. (2009). Fifteen novel mutations in PKLR associated with pyruvate kinase (PK) deficiency: structural implications of amino acid substitutions in PK. Hum Mutat *30*, 446-453

239. van Wijk, R., van Solinge, W.W., Nerlov, C., Beutler, E., Gelbart, T., Rijksen, G., and Nielsen, F.C. (2003). Disruption of a novel regulatory element in the erythroid-specific promoter of the human PKLR gene causes severe pyruvate kinase deficiency. Blood *101*, 1596-1602

240. Vigna, E., Amendola, M., Benedicenti, F., Simmons, A.D., Follenzi, A., and Naldini, L. (2005). Efficient Tet-dependent expression of human factor IX in vivo by a new self-regulating lentiviral vector. Mol Ther *11*, 763-775

241. Visigalli, I., Delai, S., Politi, L.S., Di Domenico, C., Cerri, F., Mrak, E., D'Isa, R., Ungaro, D., Stok, M., Sanvito, F.*, et al.* (2010). Gene therapy augments the efficacy of hematopoietic cell transplantation and fully corrects mucopolysaccharidosis type I phenotype in the mouse model. Blood *116*, 5130-5139

242. Vukelja, S.J. (1994). Erythropoietin in the treatment of iron overload in a patient with hemolytic anemia and pyruvate kinase deficiency. Acta Haematol *91*, 199-200

243. Walkley, C.R. (2011). Erythropoiesis, anemia and the bone marrow microenvironment. Int J Hematol *93*, 10-13

244. Wang, C., Chiarelli, L.R., Bianchi, P., Abraham, D.J., Galizzi, A., Mattevi, A., Zanella, A., and Valentini, G. (2001). Human erythrocyte pyruvate kinase: characterization of the recombinant enzyme and a mutant form (R510Q) causing nonspherocytic hemolytic anemia. Blood *98*, 3113-3120

245. Wang, G.L., and Semenza, G.L. (1996). Molecular basis of hypoxia-induced erythropoietin expression. Curr Opin Hematol *3*, 156-162

246. Whitney, K.M., Goodman, S.A., Bailey, E.M., and Lothrop, C.D., Jr. (1994). The molecular basis of canine pyruvate kinase deficiency. Exp Hematol *22*, 866-874

247. Winterbourn, C.C. (1990). Oxidative denaturation in congenital hemolytic anemias: the unstable hemoglobins. Semin Hematol *27*, 41-50

248. Wu, X., Li, Y., Crise, B., and Burgess, S.M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. Science *300*, 1749-1751

249. Wu, H., Liu, X., Jaenisch, R., and Lodish, H.F. (1995). Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. Cell *83*, 59-67

250. Yahata, T., Takanashi, T., Muguruma, Y., Ibrahim, A.A., Matsuzawa, H., Uno, T., Sheng, Y., Onizuka, M., Ito, M., Kato, S.*, et al.* (2011). Accumulation of oxidative DNA damage restricts the self-renewal capacity of human hematopoietic stem cells. Blood *118*, 2941-2950

251. Yang, L., Bailey, L., Baltimore, D., and Wang, P. (2006). Targeting lentiviral vectors to specific cell types in vivo. Proc Natl Acad Sci U S A *103*, 11479-11484

252. Yew, N.S., Zhao, H., Przybylska, M., Wu, I.H., Tousignant, J.D., Scheule, R.K., and Cheng, S.H. (2002). CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo. Mol Ther 5, 731-738

253. Zaiss, A.K., Son, S., and Chang, L.J. (2002). RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. J Virol *76*, 7209-7219

254. Zanella, A., and Bianchi, P. (2000). Red cell pyruvate kinase deficiency: from genetics to clinical manifestations. Baillieres Best Pract Res Clin Haematol *13*, 57-81

255. Zanella, A., Bianchi, P., Baronciani, L., Zappa, M., Bredi, E., Vercellati, C., Alfinito, F., Pelissero, G., and Sirchia, G. (1997). Molecular characterization of PK-LR gene in pyruvate kinase-deficient Italian patients. Blood *89*, 3847-3852

256. Zanella, A., Berzuini, A., Colombo, M.B., Guffanti, A., Lecchi, L., Poli, F., Cappellini, M.D., and Barosi, G. (1993). Iron status in red cell pyruvate kinase deficiency: study of Italian cases. Br J Haematol *83*, 485-490

257. Zanella, A., Bianchi, P., Fermo, E., Iurlo, A., Zappa, M., Vercellati, C., Boschetti, C., Baronciani, L., and Cotton, F. (2001a). Molecular characterization of the PK-LR gene in sixteen pyruvate kinase-deficient patients. Br J Haematol *113*, 43-48

258. Zanella, A., Bianchi, P., Iurlo, A., Boschetti, C., Taioli, E., Vercellati, C., Zappa, M., Fermo, E., Tavazzi, D., and Sampietro, M. (2001b). Iron status and HFE genotype in erythrocyte pyruvate kinase deficiency: study of Italian cases. Blood Cells Mol Dis *27*, 653-661

259. Zanella, A., Fermo, E., Bianchi, P., Chiarelli, L.R., and Valentini, G. (2007). Pyruvate kinase deficiency: the genotype-phenotype association. Blood Rev *21*, 217-231

260. Zanella, A., Fermo, E., Bianchi, P., and Valentini, G. (2005). Red cell pyruvate kinase deficiency: molecular and clinical aspects. Br J Haematol *130*, 11-25

261. Zanta-Boussif, M.A., Charrier, S., Brice-Ouzet, A., Martin, S., Opolon, P., Thrasher, A.J., Hope, T.J., and Galy, A. (2009). Validation of a mutated PRE sequence allowing high and sustained transgene expression while abrogating WHV-X protein synthesis: application to the gene therapy of WAS. Gene Ther *16*, 605-619

262. Zaucha, J.A., Yu, C., Lothrop, C.D., Jr., Nash, R.A., Sale, G., Georges, G., Kiem, H.P., Niemeyer, G.P., Dufresne, M., Cao, Q*., et al.* (2001). Severe canine hereditary hemolytic anemia treated by nonmyeloablative marrow transplantation. Biol Blood Marrow Transplant *7*, 14-24

263. Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., and Charneau, P. (2000). HIV-1 genome nuclear import is mediated by a central DNA flap. Cell *101*, 173-185

264. Zerez, C.R., and Tanaka, K.R. (1987). Impaired nicotinamide adenine dinucleotide synthesis in pyruvate kinase-deficient human erythrocytes: a mechanism for decreased total NAD content and a possible secondary cause of hemolysis. Blood *69*, 999-1005

265. Zhao, H., Pestina, T.I., Nasimuzzaman, M., Mehta, P., Hargrove, P.W., and Persons, D.A. (2009). Amelioration of murine beta-thalassemia through drug selection of hematopoietic stem cells transduced with a lentiviral vector encoding both gamma-globin and the MGMT drug-resistance gene. Blood *113*, 5747-5756

266. Zufferey, R., Donello, J.E., Trono, D., and Hope, T.J. (1999). Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. J Virol *73*, 2886-2892

267. Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol *72*, 9873-9880

268. Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., and Trono, D. (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol *15*, 871-875

269. Zychlinski, D., Schambach, A., Modlich, U., Maetzig, T., Meyer, J., Grassman, E., Mishra, A., and Baum, C. (2008). Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol Ther *16*, 718-725

APPENDIX 1

1. Mouse R-type Pyruvate Kinase cDNA Alignment with coRPK sequence

Sequence 1: wt cDNA from *Pklr* mouse gene (Variant 1 or R-type isoform CCDS Gene Bank NM_013631.2)
Sequence 2: synthesized coRPK
Length: 1756
Similarity: 1344/1756 (76.5%)
Gaps: 44/1756 (2.5%)

```
wt Pklr      1 ------------ATGTCTGTCCAAGAGAA---CGAGCTACCCCAGCAGCT     35
                           |||....||||.||.||    | ||||  |.|.||||||
coRPK        1 ACCGGTGCCACCATGAGCATCCAGGAAAATATC-AGCT--CTCTGCAGCT     47

wt Pklr     36 CTGGCCCTGGATCTTTAAGTCCCAAAAAGACTTAGCAAAGTCTGCTT--T     83
               ..||.||||||.|.|..|||..|||.|.||||.|.||.|||  .||.|  |
coRPK       48 GCGGTCCTGGGTGTCCAAGAGCCAGAGAGACCTGGCCAAG--AGCATCCT     95

wt Pklr     84 AAGTGGGGCTCCAGGAGGGCCAGCAGGATACCTGAGACGTGCCAGTGTGG    133
               .|..||.||.||.||.||.||||||.|||||||||||||.|.||.||.||||
coRPK       96 GATCGGAGCCCCTGGCGGACCAGCCGGATACCTGAGAAGGGCTAGCGTGG    145

wt Pklr    134 CCCAGCTGACCCAGGAGCTGGGCACTGCCTTCTTCCAGCAGCAGCAACTG    183
               |||||||||||||||.||||||||.||||||.|||||||||||||||.|||
coRPK      146 CCCAGCTGACCCAGGAACTGGGCACCGCCTTTTTCCAGCAGCAGCAGCTG    195

wt Pklr    184 CCCGCAGCTATGGCGGACACCTTCCTGGAACACCTCTGCCTTCTGGATAT    233
               ||.||.||.||||||.|||||||||.||||||||||||.||||||.||||||.||
coRPK      196 CCAGCCGCCATGGCCGACACCTTTCTGGAACACCTGTGCCTGCTGGACAT    245

wt Pklr    234 CGACTCAGAGCCTGTGGCCGCTAGGAGCACCAGCATCATTGCCACCATCG    283
               ||||||.||||||.|||||||||.||.||||||||||||||||||||||||||
coRPK      246 CGACTCTGAGCCCGTGGCCGCCAGAAGCACCAGCATCATTGCCACCATCG    295

wt Pklr    284 GGCCGGCGTC-----ACGCTCTGTGGACCGCCTCAAGGAGATGATCAAGG    328
               |.||.||  |     |.||   |||||.||.||.||.||||||||||||||
coRPK      296 GCCCTGC--CAGCAGAAGC---GTGGAGCGGCTGAAAGAGATGATCAAGG    340

wt Pklr    329 CAGGGATGAACATTGCACGACTCAACTTCTCCCATGGCTCCCATGAGTAC    378
               |.||.||||||.||.||.||.||.||||||||||||.|||..|||.||||||
coRPK      341 CCGGCATGAATATCGCCCGGCTGAACTTCTCCCACGGCAGCCACGAGTAC    390

wt Pklr    379 CATGCAGAGTCCATCGCCAACATTCGGGAGGCGGCTGAAAGCTTTGCAAC    428
               ||.|||||||..|||.|||||||.|.|||||||||.|..||.|||||||||...
coRPK      391 CACGCAGAGAGCATTGCCAACGTCCGGGAGGCCGTGGAGAGCTTTGCCGG    440

wt Pklr    429 CTCCCCACTCAGCTACAGACCCGTGGCCATCGCCCTGGACACCAAGGGTC    478
               |..|||.||.|||||||||||||||||||.|||||||||||||||||||.|
coRPK      441 CAGCCCCCTGAGCTACAGACCCGTGGCCATTGCCCTGGACACCAAGGGCC    490

wt Pklr    479 CCGAGATACGCACTGGAGTCCTGCAGGGGGGTCCAGAGTCGGAGGTGGAA    528
               |||||||..|.||.|||.|.||||||||.||.||.|||...||||||||.
coRPK      491 CCGAGATCAGAACAGGAATTCTGCAGGGAGGGCCTGAGAGCGAGGTGGAG    540

wt Pklr    529 ATTGTGAAGGGTTCA--CAGGTGCTGGTGACTGTGGATCCGAAGTTCCGG    576
               .|.||||||||  ||  ||.||||||||||||.|||||.||....|||.|.
coRPK      541 CTGGTGAAGGG--CAGCCAAGTGCTGGTGACCGTGGACCCCGCCTTCAGA    588

wt Pklr    577 ACAAGGGGCGATGCAAAGACAGTGTGGGTGGACTACCACAATATCACCCA    626
               ||.||.|||.|.||.||.||||||||||||||||||.|||.|||...|.
coRPK      589 ACCAGAGGCAACGCCAACACAGTGTGGGTGGACTACCCCAACATCGTGCG    638
```

```
wt Pklr          627 GGTCGTTGCAGTGGGGGGCCGCATCTACATTGACGACGGGCTCATCTCCT  676
                     |||.||..|.|||||.|||.|.||||||||.||||||||.||.|||..|.
coRPK            639 GGTGGTGCCTGTGGGCGGCAGAATCTACATCGACGACGGCCTGATCAGCC  688

wt Pklr          677 TAGTGGTGCGGAAAATTGGCCCAGAGGGACTGGTGACCGAAGTGGAACAC  726
                     |.|||||||.|||.||.||.||.|||||.|||||||||.|.||.||..|.
coRPK            689 TGGTGGTGCAGAAGATCGGACCTGAGGGCCTGGTGACCCAGGTCGAGAAT  738

wt Pklr          727 GGTGGTTTCTTGGGCAACAGGAAGGGTGTGAACTTGCCAAATGCCGAGGT  776
                     ||.||..|..||||||.|||.|||||.||||||..|||||...|||.||||
coRPK            739 GGCGGCGTGCTGGGCAGCAGAAAGGGCGTGAATCTGCCAGGCGCCCAGGT  788

wt Pklr          777 GGACCTGCCTGGGCTATCAGAGCAAGATCTTTTG-GATCTGCGCTTCGGG  825
                     ||||||||||||.||.||.|||||.|| |.|..| ||.|||.|.||.||.
coRPK            789 GGACCTGCCTGGCCTGTCTGAGCAGGA-CGTGAGAGACCTGAGATTTGGC  837

wt Pklr          826 GTGGAGCATAATGTGGACATCATCTTTGCCTCCTTTGTACGAAAAGCCAG  875
                     ||||||||....|||||||||.|.||.|||..|||.||.||.||.|||..
coRPK            838 GTGGAGCACGGCGTGGACATCGTGTTCGCCAGCTTCGTGCGGAAGGCCTC  887

wt Pklr          876 TGATGTGGTGGCAGTCCGAGATGCCCTAGGGCCAGAAGGACGGGGCATCA  925
                     ||||||||..||.||..|||..||.||.||.||.|||||.|..|||||||
coRPK            888 TGATGTGGCCGCCGTGAGAGCCGCTCTGGGCCCTGAAGGCCACGGCATCA  937

wt Pklr          926 AAATTATCAGCAAAATCGAGAACCATGAAGGCGTGAAGAAGTTTGATGAG  975
                     |.||.||||||||.||||||||||.||.|||||||||..|||.||.|||
coRPK            938 AGATCATCAGCAAGATCGAGAACCACGAGGGCGTGAAGCGGTTCGACGAG  987

wt Pklr          976 ATCCTAGAAGTGAGCGATGGCATCATGGTGGCTCGGGGTGACCTTGGCAT 1025
                     |||||.|||||||..|||.||||||||||||||..|.||.||||||.|||||
coRPK            988 ATCCTGGAAGTGTCCGACGGCATCATGGTGGCCAGAGGCGACCTGGGCAT 1037

wt Pklr          1026 TGAGATCCCAGCAGAGAAGGTTTTCTTGGCTCAGAAGATGATGATTGGAC 1075
                      .||||||||.||.||||||||.|||.|||||.|||||.||||||||.||||
coRPK            1038 CGAGATCCCCGCCGAGAAGGTGTTCCTGGCCCAGAAAATGATGATCGGAC 1087

wt Pklr          1076 GCTGCAACCTGGCTGGCAAGCCTGTCGTTTGTGCCACACAGATGCTGGAG 1125
                      |.||||||||||.|||||.|||||.||.||.|||||.||||||||||||.
coRPK            1088 GGTGCAACCTGGCCGGCAAACCTGTGGTGTGCGCCACCCAGATGCTGGAA 1137

wt Pklr          1126 AGCATGATCACTAAGGCTCGACCAACTCGGGCGGAGACAAGCGATGTGGC 1175
                      ||||||||||.|||.|..|||||.||..|.|||.||||||||||||.|||||
coRPK            1138 AGCATGATCACCAAGCCCAGACCCACCAGAGCCGAGACAAGCGACGTGGC 1187

wt Pklr          1176 CAATGCTGTGCTGGATGGGGCTGACTGTATCATGCTGTCTGGAGAGACCG 1225
                      |||.||.||||||||||.|||||||.||||||||||||.||.|||||.|
coRPK            1188 CAACGCCGTGCTGGATGGCGCTGACTGCATCATGCTGTCCGGCGAGACAG 1237

wt Pklr          1226 CCAAAGGCAGTTTCCCCGTGGAGGCTGTAAAGATGCAACATGCGATTGCC 1275
                      ||||.||||..|||||||||||||.||.||||||||.||.||.||||||
coRPK            1238 CCAAGGGCAACTTCCCCGTGGAGGCCGTGAAGATGCAGCACGCCATTGCC 1287

wt Pklr          1276 CGGGAGGCAGAGGCCGCTGTGTACCACCGCCAGTTGTTTGAGGAGCTACG 1325
                      .|.||.||.||||||||.||||||||||.||||.||||.||||||.||.||
coRPK            1288 AGAGAAGCCGAGGCCGCCGTGTACCACCGGCAGCTGTTCGAGGAACTGCG 1337

wt Pklr          1326 CCGGGCAGCGCCGCTGAGCCGTGACCCAACCGAGGTCACTGCAATTGGAG 1375
                      ..|.||.||.||.||||||.|.||.||.||||||.||.||.||.||.||||
coRPK            1338 GAGAGCCGCCCCTCTGAGCAGAGATCCCACCGAAGTGACCGCCATCGGAG 1387

wt Pklr          1376 CCGTGGAGGCTTCCTTCAAGTGCTGTGCCGCAGCCATCATTGTGCTGACA 1425
                      |||||||.||..||||||||||||||.||||||.||.||||||.||||||||.
coRPK            1388 CCGTGGAAGCCGCCTTCAAGTGCTGCGCCGCTGCAATCATCGTGCTGACC 1437

wt Pklr          1426 AAGACTGGCCGCTCAGCTCAGCTTCTGTCTCGCTACCGACCTCGGGCTGC 1475
                      |..||.||||.|....||.||||||.||||||..|.||||.||||..|.||.||
coRPK            1438 ACCACAGGCAGAAGCGCCCAGCTGCTGTCCAGATACAGACCCAGAGCCGC 1487

wt Pklr          1476 TGTCATTGCTGTGACTCGTTCTGCCCAGGCTGCCCGACAGGTCCACCTGT 1525
                      .||.||.||.||||||..|.||.||||||||.||..||||||||||||||||
coRPK            1488 CGTGATCGCCGTGACAAGATCCGCCCAGGCCGCTAGACAGGTCCACCTGT 1537

wt Pklr          1526 CCCGAGGAGTCTTCCCCTTGCTCTACCGTGAGCCTCCAGAGGCTGTCTGG 1575
                      .|.||||.||.||||||.||||.||||||.||||||||.||||||..||||||
coRPK            1538 GCAGAGGCGTGTTCCCCCTGCTGTACCGGGAGCCTCCCGAGGCCATCTGG 1587

wt Pklr          1576 GCAGATGATGTGGACCGAAGGGTCCAATTTGGCATCGAAAGTGGAAAGCT 1625
                      ||.||.||.||||||.||.|||||.||.||.||||||||.||.||.||||||
coRPK            1588 GCCGACGACGTGGACAGACGGGTGCAGTTCGGCATCGAGAGCGGCAAGCT 1637
```

```
wt Pklr          1626 TCGTGGCTTCCTCCGAGTTGGTGATCTGGTGATTGTGGTGACAGGCTGGC   1675
                      .||.|||||||||..||||.||.||.|||||||||.||||||||||||||||
coRPK            1638 GCGGGGCTTCCTGAGAGTGGGCGACCTGGTGATCGTGGTGACAGGCTGGC   1687

wt Pklr          1676 GACCTGGCTCTGGCTATACCAACATTATGCGGGTGCTGACCATATCCTGA   1725
                      |.|||||||...|||||.|||||||||.|||.|||||||||.||||...||||
coRPK            1688 GGCCTGGCAGCGGCTACACCAACATCATGAGGGTGCTGTCCATCAGCTGA   1737

wt Pklr          1726 ------   1725

coRPK            1738 CCGCGG   1743
```

2. Human R-type pyruvate kinase cDNA alignment with coRPK sequence

Sequence 1: wt cDNA from *PKLR* human gene (Variant 1 or R-type isoform, CCDS Gene Bank NM_000298.5)
Sequence 2: synthesized coRPK
Length: 1749
Similarity: 1407/1749 (80.4%)
Gaps: 30/1749 (1.7%)

```
wt PKLR             1 ------------ATGTCGATCCAGGAGAACATATCA--TCCCTGCAGCTT     36
                                  |||...||||||  |||.||||||  ||.|||||||||.
coRPK               1 ACCGGTGCCACCATGAGCATCCAG--GAAAATATCAGCTCTCTGCAGCTG     48

wt PKLR            37 CGGTCATGGGTCTCTAAGTCCCAAAGAGACTTAGCAAAGTCCATCCTGAT     86
                      |||||.||||||.||.|||..|||.|||||||.|.||.|||..|||||||||
coRPK              49 CGGTCCTGGGTGTCCAAGAGCCAGAGAGACCTGGCCAAGAGCATCCTGAT     98

wt PKLR            87 TGGGGCTCCAGGAGGGCCAGCGGGGTATCTGCGGCGGGCCAGTGTGGCCC    136
                      .||.||.||.||.||.||||||.||.||.|||.|..||||.||.|||||||
coRPK              99 CGGAGCCCCTGGCGGACCAGCCGGATACCTGAGAAGGGCTAGCGTGGCCC    148

wt PKLR           137 AACTGACCCAGGAGCTGGGCACTGCCTTCTTCCAGCAGCAGCAGCTGCCA    186
                      |.||||||||||.|||||||.||||||.||||||||||||||||||||||||
coRPK             149 AGCTGACCCAGGAACTGGGCACCGCCTTTTTCCAGCAGCAGCAGCTGCCA    198

wt PKLR           187 GCTGCTATGGCAGACACCTTCCTGGAACACCTCTGCCTACTGGACATTGA    236
                      ||.||.||||||.|||||||.||||||||||.||||||.|||||||||.||
coRPK             199 GCCGCCATGGCCGACACCTTTCTGGAACACCTGTGCCTGCTGGACATCGA    248

wt PKLR           237 CTCCGAGCCCGTGGCTGCTCGCAGTACCAGCATCATTGCCACCATCGG--    284
                      |||.||||||||||||.||..|.||.||||||||||||||||||||||
coRPK             249 CTCTGAGCCCGTGGCCGCCAGAAGCACCAGCATCATTGCCACCATCGGCC    298

wt PKLR           285 --GCCAGCATCTCGCTCCGTGGAGCGCCTCAAGGAGATGATCAAGGCCGG    332
                        |||||||    |...||||||||||.||.||.||||||||||||||||
coRPK             299 CTGCCAGCA----GAAGCGTGGAGCGGCTGAAAGAGATGATCAAGGCCGG    344

wt PKLR           333 GATGAACATTGCGCGACTCAACTTCTCCCACGGCTCCCACGAGTACCATG    382
                      .||||||.||.||.||.||||||||||||||..||||||||||||||.|
coRPK             345 CATGAATATCGCCCGGCTGAACTTCTCCCACGGCAGCCACGAGTACCACG    394

wt PKLR           383 CTGAGTCCATCGCCAACGTCCGGGAGGCGGTGGAGAGCTTTGCAGGTTCC    432
                      |.|||..|||.||||||||||||||||.|||||||||||||.||...|
coRPK             395 CAGAGAGCATTGCCAACGTCCGGGAGGCCGTGGAGAGCTTTGCCGGCAGC    444
```

```
wt PKLR          433 CCACTCAGCTACCGGCCCGTGGCCATCGCCCTGGACACCAAGGGACCGGA  482
                     ||.||.|||||||.|.||||||||||.||||||||||||||||.||.||
coRPK            445 CCCCTGAGCTACAGACCCGTGGCCATTGCCCTGGACACCAAGGGCCCCGA  494

wt PKLR          483 GATCCGCACTGGGATCCTGCAGGGGGGTCCAGAGTCGGAAGTGGAGCTGG  532
                     ||||.|.||.||.||.||||||||.||.||.|||...||.||||||||||
coRPK            495 GATCAGAACAGGAATTCTGCAGGGAGGGCCTGAGAGCGAGGTGGAGCTGG  544

wt PKLR          533 TGAAGGGCTCCCAGGTGCTGGTGACTGTGGACCCCGCGTTCCGGACGCGG  582
                     ||||||||..|||.||||||||||.||||||||||.|||.|.||..|.
coRPK            545 TGAAGGGCAGCCAAGTGCTGGTGACCGTGGACCCCGCCTTCAGAACCAGA  594

wt PKLR          583 GGGAACGCGAACACCGTGTGGGTGGACTACCCCAATATTGTCCGGGTCGT  632
                     ||.|||||.|||||.||||||||||||||||.||.||.||||||.||
coRPK            595 GGCAACGCCAACACAGTGTGGGTGGACTACCCCAACATCGTGCGGGTGGT  644

wt PKLR          633 GCCGGTGGGGGGCCGCATCTACATTGACGACGGGCTCATCTCCCTAGTGG  682
                     |||.|||||.|||.|.||||||||.||||||||.||.|||..|||.||||
coRPK            645 GCCTGTGGGCGGCAGAATCTACATCGACGACGGCCTGATCAGCCTGGTGG  694

wt PKLR          683 TCCAGAAAATCGGCCCAGAGGGACTGGTGACCCAAGTGGAGAACGGCGGC  732
                     |.|||||.|||||.||.||||||.||||||||||.||.||||||.||||||
coRPK            695 TGCAGAAGATCGGACCTGAGGGCCTGGTGACCCAGGTCGAGAATGGCGGC  744

wt PKLR          733 GTCCTGGGCAGCCGGAAGGGCGTGAACTTGCCAGGGGCCCAGGTGGACTT  782
                     ||.||||||||||.|.||||||||||..||||||||.||||||||||||.|
coRPK            745 GTGCTGGGCAGCAGAAAGGGCGTGAATCTGCCAGGCGCCCAGGTGGACCT  794

wt PKLR          783 GCCCGGGCTGTCCGAGCAGGACGTCCGAGACCTGCGCTTCGGGGTGGAGC  832
                     |||.||.||||||.||||||||||||..|||||||||.|.||.||.|||||||
coRPK            795 GCCTGGCCTGTCTGAGCAGGACGTGAGAGACCTGAGATTTGGCGTGGAGC  844

wt PKLR          833 ATGGGGTGGACATCGTCTTTGCCTCCTTTGTGCGGAAAGCCAGCGACGTG  882
                     |.||.||||||||||.||.||||..|||.||||||||.|||...||.|||
coRPK            845 ACGGCGTGGACATCGTGTTCGCCAGCTTCGTGCGGAAGGCCTCTGATGTG  894

wt PKLR          883 GCTGCCGTCAGGGCTGCTCTGGGTCCGGAAGGACACGGCATCAAGATCAT  932
                     ||.||||||.||.||.|||||||||.||.||||||.||||||||||||||||
coRPK            895 GCCGCCGTGAGAGCCGCTCTGGGCCCTGAAGGCCACGGCATCAAGATCAT  944

wt PKLR          933 CAGCAAAATTGAGAACCACGAAGGCGTGAAGAGGTTTGATGAAATCCTGG  982
                     ||||||.||.||||||||||.|||||||||.|||||.||.||.|||||||
coRPK            945 CAGCAAGATCGAGAACCACGAGGGCGTGAAGCGGTTCGACGAGATCCTGG  994

wt PKLR          983 AGGTGAGCGACGGCATCATGGTGGCACGGGGGGACCTAGGCATCGAGATC 1032
                     |.|||..||||||||||||||||||..|.||.||||||.||||||||||||
coRPK            995 AAGTGTCCGACGGCATCATGGTGGCCAGAGGCGACCTGGGCATCGAGATC 1044

wt PKLR         1033 CCAGCAGAGAAGGTTTTCCTGGCTCAGAAGATGATGATTGGGCGCTGCAA 1082
                     ||.||.||||||||.||||||||.|||||.||||||||.||.||.|||||
coRPK           1045 CCCGCCGAGAAGGTGTTCCTGGCCCAGAAAATGATGATCGGACGGTGCAA 1094

wt PKLR         1083 CTTGGCGGGCAAGCCTGTTGTCTGTGCCACACAGATGCTGGAGAGCATGA 1132
                     |.||||.||||||.|||||.||.||.|||||.||||||||||.|||||||
coRPK           1095 CCTGGCCGGCAAACCTGTGGTGTGCGCCACCCAGATGCTGGAAAGCATGA 1144

wt PKLR         1133 TTACCAAGCCCCGGCCAACGAGGGCAGAGACAAGCGATGTCGCCAATGCT 1182
                      |.|||||||||.|.||.||.||.||.||||||||||.||.||||||.||.
coRPK           1145 TCACCAAGCCCAGACCCACCAGAGCCGAGACAAGCGACGTGGCCAACGCC 1194

wt PKLR         1183 GTGCTGGATGGGGCTGACTGCATCATGCTGTCAGGGGAGACTGCCAAGGG 1232
                     ||||||||||.|||||||||||||||||||.||.|||||.||||||||
coRPK           1195 GTGCTGGATGGCGCTGACTGCATCATGCTGTCCGGCGAGACAGCCAAGGG 1244

wt PKLR         1233 CAACTTCCCTGTGGAAGCGGTGAAGATGCAGCATGCGATTGCCCGGGAGG 1282
                     |||||||||.|||||.||.||||||||||||||.||.||||||.|.||.|
coRPK           1245 CAACTTCCCCGTGGAGGCCGTGAAGATGCAGCACGCCATTGCCAGAGAAG 1294

wt PKLR         1283 CAGAGGCCGCAGTGTACCACCGGCAGCTGTTTGAGGAGCTACGTCGGGCA 1332
                     |.||||||||.|||||||||||||||||||.||||||.||.||..|.||.
coRPK           1295 CCGAGGCCGCCGTGTACCACCGGCAGCTGTTCGAGGAACTGCGGAGAGCC 1344

wt PKLR         1333 GCGCCACTAAGCCGTGATCCCACTGAGGTCACCGCCATTGGTGCTGTGGA 1382
                        ||.||.||.|||.|.||||||||.||.||.||||||||.||.||.|||||
coRPK           1345 GCCCCTCTGAGCAGAGATCCCACCGAAGTGACCGCCATCGGAGCCGTGGA 1394

wt PKLR         1383 GGCTGCCTTCAAGTGCTGTGCTGCTGCCATCATTGTGCTGACCACAACTG 1432
                     .||.|||||||||||||.||.||||||.||||||.||||||||||.||.|
coRPK           1395 AGCCGCCTTCAAGTGCTGCGCCGCTGCAATCATCGTGCTGACCACCACAG 1444
```

```
wt PKLR          1433 GCCGCTCAGCCCAGCTTCTGTCTCGGTACCGACCTCGGGCAGCAGTCATT   1482
                      ||.|....|||||||.|||||..|.|||.||||..|.||.||.||.||.
coRPK            1445 GCAGAAGCGCCCAGCTGCTGTCCAGATACAGACCCAGAGCCGCCGTGATC   1494

wt PKLR          1483 GCTGTCACCCGCTCTGCCCAGGCTGCCCGCCAGGTCCACTTATGCCGAGG   1532
                      ||.||.||..|.||.|||||||||.||..|.|||||||||.|.|||.||||
coRPK            1495 GCCGTGACAAGATCCGCCCAGGCCGCTAGACAGGTCCACCTGTGCAGAGG   1544

wt PKLR          1533 AGTCTTCCCCTTGCTTTACCGTGAACCTCCAGAAGCCATCTGGGCAGATG   1582
                      .||.||||||.||||.|||||.||.||||||.||.||||||||||||.||.|
coRPK            1545 CGTGTTCCCCCTGCTGTACCGGGAGCCTCCCGAGGCCATCTGGGCCGACG   1594

wt PKLR          1583 ATGTAGATCGCCGGGTGCAATTTGGCATTGAAAGTGGAAAGCTCCGTGGC   1632
                      |.||.||..|.|||||||.||.|||||.||.||.||.|||||.||.|||
coRPK            1595 ACGTGGACAGACGGGTGCAGTTCGGCATCGAGAGCGGCAAGCTGCGGGGC   1644

wt PKLR          1633 TTCCTCCGTGTTGGAGACCTGGTGATTGTGGTGACAGGCTGGCGACCTGG   1682
                      ||||||..|.||.||.||||||||||.|||||||||||||||||.|||||
coRPK            1645 TTCCTGAGAGTGGGCGACCTGGTGATCGTGGTGACAGGCTGGCGGCCTGG   1694

wt PKLR          1683 CTCCGGCTACACCAACATCATGCGGGTGCTAAGCATATCCTGA------   1725
                      |..|||||||||||||||||.|||||||...|||...||||
coRPK            1695 CAGCGGCTACACCAACATCATGAGGGTGCTGTCCATCAGCTGACCGCGG   1743
```

3. Mouse R-type Pyruvate Kinase and coRPK Protein Alignment

Sequence 1: wt mouse RPK (Variant 1 or R-type isoform)
Sequence 2: coRPK protein
Length: 578
Similarity: 548/578 (94.8%)
Gaps: 4/578 (0.7%)

```
wt Pklr             1 ----MSVQENELPQQLWPWIFKSQKDLAKSALSGAPGGPAGYLRRASVAQ     46
                        ||:|||....||..|:.|||:|||||.|.|||||||||||||||||||
coRPK               1 TGATMSIQENISSLQLRSWVSKSQRDLAKSILIGAPGGPAGYLRRASVAQ     50

wt Pklr            47 LTQELGTAFFQQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIIATIGP     96
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK              51 LTQELGTAFFQQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIIATIGP    100

wt Pklr            97 ASRSVDRLKEMIKAGMNIARLNFSHGSHEYHAESIANIREAAESFATSPL    146
                      |||||:|||||||||||||||||||||||||||||:|||.|||||.|||
coRPK             101 ASRSVERLKEMIKAGMNIARLNFSHGSHEYHAESIANVREAVESFAGSPL    150

wt Pklr           147 SYRPVAIALDTKGPEIRTGVLQGGPESEVEIVKGSQVLVTVDPKFRTRGD    196
                      ||||||||||||||||||:||||||||||:||||||||||.||||||:
coRPK             151 SYRPVAIALDTKGPEIRTGILQGGPESEVELVKGSQVLVTVDPAFRTRGN    200

wt Pklr           197 AKTVWVDYHNITQVVAVGGRIYIDDGLISLVVRKIGPEGLVTEVEHGGFL    246
                      |.||||||.||.:||.|||||||||||||||:|||||||||:||:||.|
coRPK             201 ANTVWVDYPNIVRVVPVGGRIYIDDGLISLVVQKIGPEGLVTQVENGGVL    250

wt Pklr           247 GNRKGVNLPNAEVDLPGLSEQDLLDLRFGVEHNVDIIFASFVRKASDVVA    296
                      |:|||||||.|:||||||||||:.||||||||.|||:||||||||||.|
coRPK             251 GSRKGVNLPGAQVDLPGLSEQDVRDLRFGVEHGVDIVFASFVRKASDVAA    300

wt Pklr           297 VRDALGPEGRGIKIISKIENHEGVKKFDEILEVSDGIMVARGDLGIEIPA    346
                      ||.|||||||||||||||||:||||||||||||||||||||||||||||
coRPK             301 VRAALGPEGHGIKIISKIENHEGVKRFDEILEVSDGIMVARGDLGIEIPA    350
```

```
wt Pklr          347 EKVFLAQKMMIGRCNLAGKPVVCATQMLESMITKARPTRAETSDVANAVL    396
                     ||||||||||||||||||||||||||||||||.|||||||||||||||
coRPK            351 EKVFLAQKMMIGRCNLAGKPVVCATQMLESMITKPRPTRAETSDVANAVL    400

wt Pklr          397 DGADCIMLSGETAKGSFPVEAVKMQHAIAREAEAAVYHRQLFEELRRAAP    446
                     ||||||||||||||:||||||||||||||||||||||||||||||||||
coRPK            401 DGADCIMLSGETAKGNFPVEAVKMQHAIAREAEAAVYHRQLFEELRRAAP    450

wt Pklr          447 LSRDPTEVTAIGAVEASFKCCAAAIIVLTKTGRSAQLLSRYRPRAAVIAV    496
                     ||||||||||||||||:||||||||||||.||||||||||||||||||||
coRPK            451 LSRDPTEVTAIGAVEAAFKCCAAAIIVLTTTGRSAQLLSRYRPRAAVIAV    500

wt Pklr          497 TRSAQAARQVHLSRGVFPLLYREPPEAVWADDVDRRVQFGIESGKLRGFL    546
                     ||||||||||||.|||||||||||||||:||||||||||||||||||||
coRPK            501 TRSAQAARQVHLCRGVFPLLYREPPEAIWADDVDRRVQFGIESGKLRGFL    550

wt Pklr          547 RVGDLVIVVTGWRPGSGYTNIMRVLTIS    574
                     ||||||||||||||||||||||||||:||
coRPK            551 RVGDLVIVVTGWRPGSGYTNIMRVLSIS    578
```

4. Human R-type Pyruvate Kinase and coRPK Protein Alignment

Sequence 1: wt human RPK (Variant 1 or R-type isoform)
Sequence 2: coRPK protein
Length: 578
Similarity: 574/578 (99.3%)
Gaps: 4/578 (0.7%)

```
wt PKLR            1 ----MSIQENISSLQLRSWVSKSQRDLAKSILIGAPGGPAGYLRRASVAQ     46
                       ||||||||||||||||||||||||||||||||||||||||||||||||
coRPK              1 TGATMSIQENISSLQLRSWVSKSQRDLAKSILIGAPGGPAGYLRRASVAQ     50

wt PKLR           47 LTQELGTAFFQQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIIATIGP     96
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK             51 LTQELGTAFFQQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIIATIGP    100

wt PKLR           97 ASRSVERLKEMIKAGMNIARLNFSHGSHEYHAESIANVREAVESFAGSPL    146
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            101 ASRSVERLKEMIKAGMNIARLNFSHGSHEYHAESIANVREAVESFAGSPL    150

wt PKLR          147 SYRPVAIALDTKGPEIRTGILQGGPESEVELVKGSQVLVTVDPAFRTRGN    196
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            151 SYRPVAIALDTKGPEIRTGILQGGPESEVELVKGSQVLVTVDPAFRTRGN    200

wt PKLR          197 ANTVWVDYPNIVRVVPVGGRIYIDDGLISLVVQKIGPEGLVTQVENGGVL    246
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            201 ANTVWVDYPNIVRVVPVGGRIYIDDGLISLVVQKIGPEGLVTQVENGGVL    250

wt PKLR          247 GSRKGVNLPGAQVDLPGLSEQDVRDLRFGVEHGVDIVFASFVRKASDVAA    296
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            251 GSRKGVNLPGAQVDLPGLSEQDVRDLRFGVEHGVDIVFASFVRKASDVAA    300

wt PKLR          297 VRAALGPEGHGIKIISKIENHEGVKRFDEILEVSDGIMVARGDLGIEIPA    346
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            301 VRAALGPEGHGIKIISKIENHEGVKRFDEILEVSDGIMVARGDLGIEIPA    350


wt PKLR          347 EKVFLAQKMMIGRCNLAGKPVVCATQMLESMITKPRPTRAETSDVANAVL    396
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            351 EKVFLAQKMMIGRCNLAGKPVVCATQMLESMITKPRPTRAETSDVANAVL    400
```

```
wt PKLR          397 DGADCIMLSGETAKGNFPVEAVKMQHAIAREAEAAVYHRQLFEELRRAAP    446
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            401 DGADCIMLSGETAKGNFPVEAVKMQHAIAREAEAAVYHRQLFEELRRAAP    450

wt PKLR          447 LSRDPTEVTAIGAVEAAFKCCAAAIIVLTTTGRSAQLLSRYRPRAAVIAV    496
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            451 LSRDPTEVTAIGAVEAAFKCCAAAIIVLTTTGRSAQLLSRYRPRAAVIAV    500

wt PKLR          497 TRSAQAARQVHLCRGVFPLLYREPPEAIWADDVDRRVQFGIESGKLRGFL    546
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
coRPK            501 TRSAQAARQVHLCRGVFPLLYREPPEAIWADDVDRRVQFGIESGKLRGFL    550

wt PKLR          547 RVGDLVIVVTGWRPGSGYTNIMRVLSIS    574
                     ||||||||||||||||||||||||||||
coRPK            551 RVGDLVIVVTGWRPGSGYTNIMRVLSIS    578
```

APPENDIX 2

1. Leukaemia development in mice transplanted with coRPK-transduced progenitors

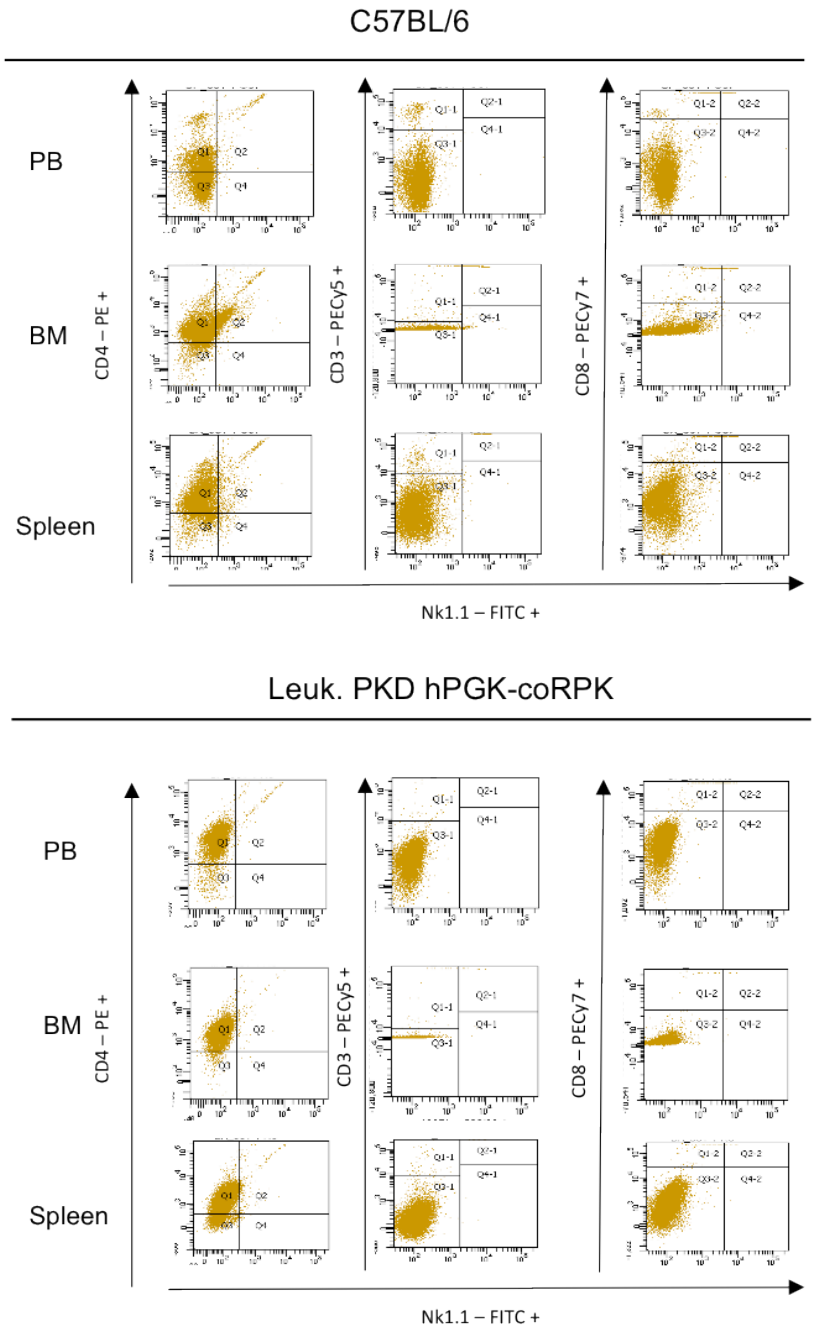

Figure 61: Study of the lymphoid compartment of leukemic mice transplanted with coRPK-transduced progenitors. Lysed samples from peripheral blood, bone marrow and spleen were labeled with a cocktail of antibodies against lymphoid markers, containing anti-mouse FITC-Nk1.1, PE-CD4, PECy5-CD3 and Biotin-CD8 antibodies, and stained with SAV-PECy7 secondary antibody. Upper part of the figure shows the lymphoid characterization of the C57BL/6 mouse used as control.

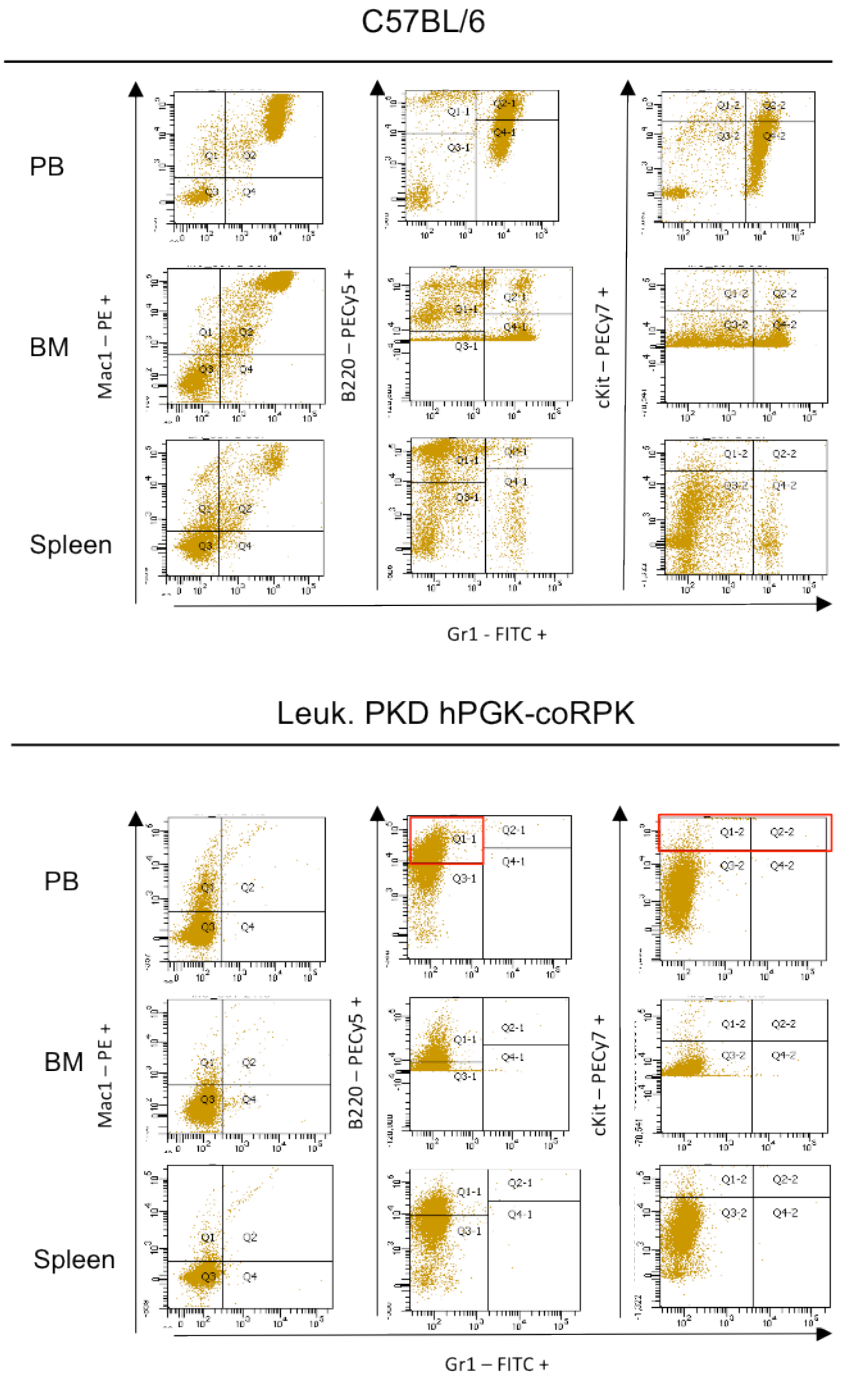

Figure 62: Study of the myeloid compartment of leukemic mice transplanted with coRPK-transduced progenitors. Lysed samples from peripheral blood, bone marrow and spleen were labeled with a cocktail of antibodies against myeloid markers, containing anti-mouse FITC-Gr1, PE-Mac1, PECy5-B220 and PECy7-cKit. Upper part of the figure shows the myeloid characterization of the C57BL/6 mouse used as control. Transplanted mice showed a B-leukaemia with no expression of c-Kit.

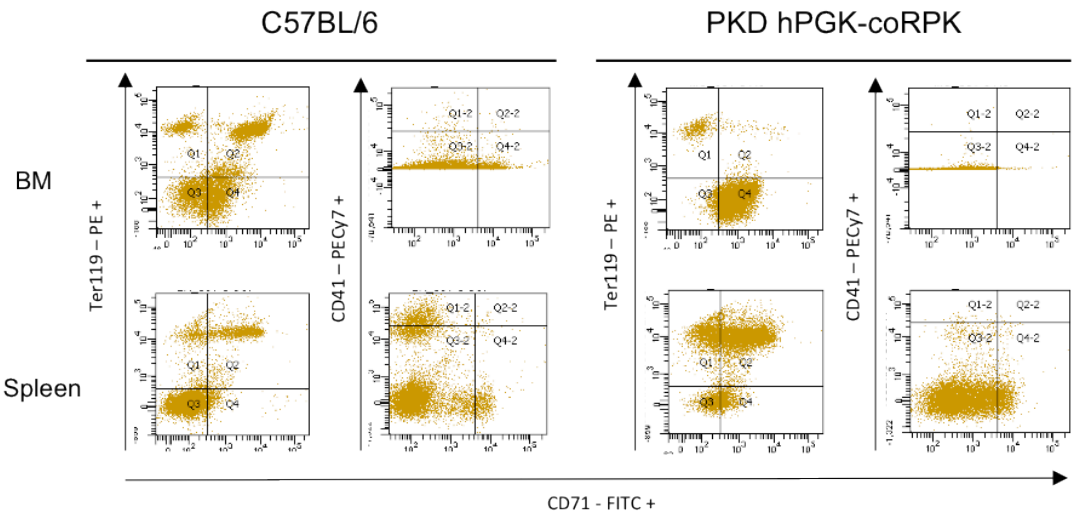

Figure 63: Study of the erythroid compartment of leukemic mice transplanted with coRPK-transduced progenitors. Non-lysed bone marrow and spleen samples were stained with PE-Ter119, FITC-CD71 and PECy7-CD41 antibodies. Left part of the figure shows the erythroid characterization of the C57BL/6 mouse used as control.

2. Leukaemia development in mice transplanted with EGFP-transduced progenitors

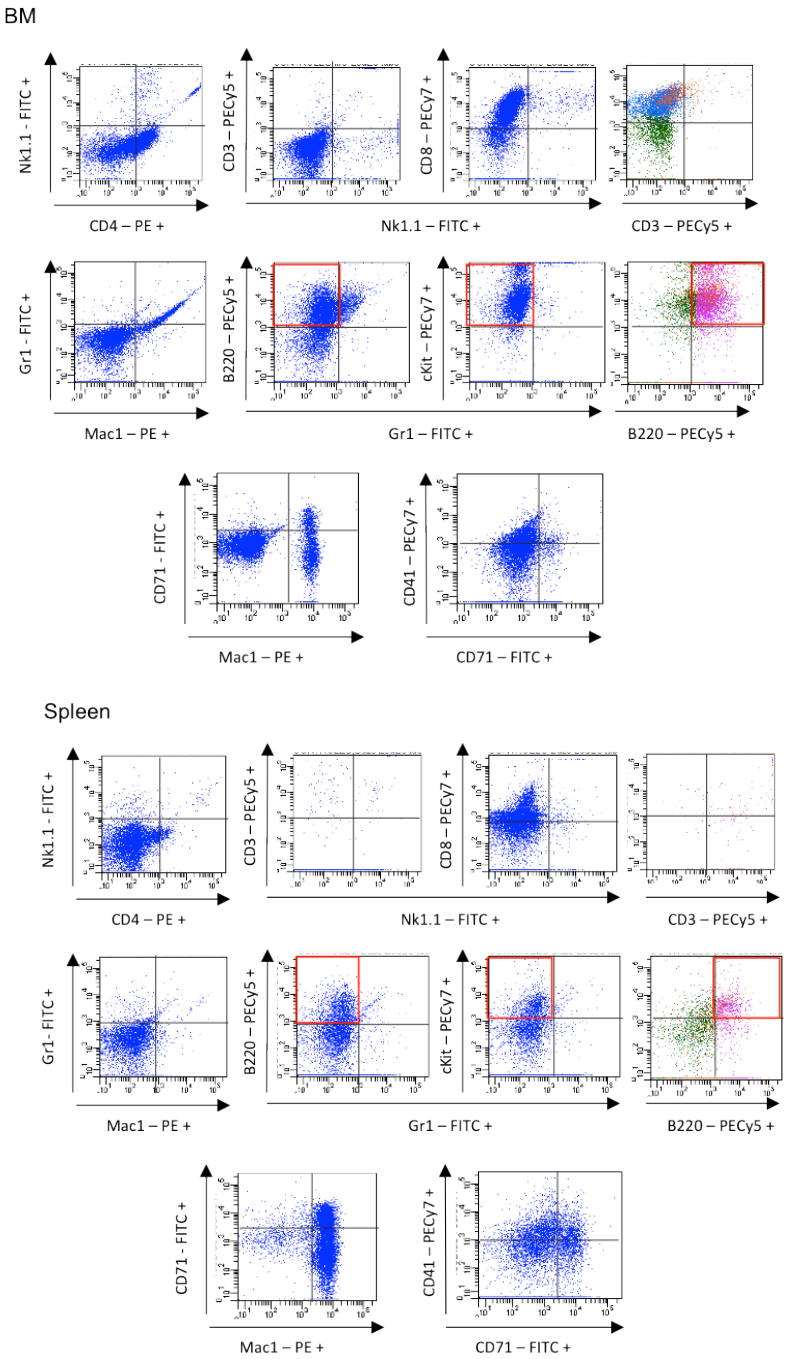

Figure 64: Study of the lymphoid and myeloid compartment of leukemic mice transplanted with EGFP-transduced progenitors. Lysed samples from bone marrow and spleen were labeled with anti-mouse FITC-Nk1.1, PE-CD4, PECy5-CD3 and Biotin-CD8 antibodies for the lymphoid characterization, and with FITC-Gr1, PE-Mac1, PECy5-B220 and PECy7-cKit antibodies for the study of myeloid markers. Transplanted mice showed a B-leukaemia with expression of c-Kit.

APPENDIX 3

Chapter 21

Gene Therapy for Erythroid Metabolic Inherited Diseases

Maria Garcia-Gomez, Oscar Quintana-Bustamante, Maria Garcia-Bravo, S. Navarro, Zita Garate and Jose C. Segovia

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/52531

1. Introduction

Gene therapy is becoming a powerful tool to treat genetic diseases. Clinical trials performed during last two decades have demonstrated its usefulness in the treatment of several genetic diseases [1] but also the need to improve vector delivery, expression and safety [2]. New vectors should reduce genotoxicity (genomic alteration due to vector integration), immunogenicity (immune response to gene delivery vectors and/or trangenes) and cytotoxicity (induced by ectopic expression and/or overexpression of the transgene).

In mature erythrocytes, most metabolic needs are covered by glycolysis, oxidative pentose phosphate pathway and glutathione cycle. Hereditary enzyme deficiencies of all these pathways have been identified, being most of them associated with chronic non-spherocytic hemolitic anemia (CNSHA). Hereditary hemolytic anemia exhibits a high molecular heterogeneity with a wide number of different mutations involved in the structural genes of nearly all affected enzymes. Deficiency in metabolic enzymes impairs energy balance in the erythrocytes, with or without changes in oxygen affinity of hemoglobin and delivery to the tissues. Despite of having a better understanding of their molecular basis, definitive curative therapy for Red Blood Cells (RBC) enzyme defects still remains undeveloped.

Conventional bone marrow transplantation allows the generation of donor-derived functional hematopoietic cells of all lineages in the host, and represents the standard of care or at least a valid therapeutic option for many inherited diseases [3]. However, complications associated to allogeneic transplantation can be as severe as the enzymatic deficiency. The recessive inheriting trait of most of these metabolic diseases and the confined enzymatic defect to the hematopoietic/

erythropoietic system, make them suitable diseases to be treated by gene therapy. Correction by gene therapy requires the stable transfer of a functional gene into the autologous self-renewing Hematopoietic stem cells (HSCs) and their mature progeny. Autologous BM transplantation of genetically corrected cells shows several advantages over the allogeneic procedure. First, it overcomes the limitation of human leukocyte antigen (HLA)-compatible donor availability, so it can be applied to every patient. Second, the reduction of morbidity and mortality associated with the transplant procedure, as there is no risk of graft versus host disease (GvHD) and consequently no need for post-transplant immunosuppression.

To date, gene therapy approaches for the treatment of inherited metabolic deficiencies are still limited, mainly because of the frequent lack of selective advantage of genetically corrected cells. This implies that high levels of transgene expression are required, as well as an efficient transduction of HSCs. This requirement have already been described in different RBC diseases as in the erytropoietic protoporphyria (EPP) [4] caused by the deficiency of the last enzyme of the heme biosintesis pathway or in the piruvate kinase deficiency (PKD) [3], where there is an impairment in the final yield of ATP in RBC. Additionally, some RBC pathologies require switching on expression of the transgene at only the proper stage of differentiation, which represents another challenge in the development of new gene therapy protocols.

2. Gene therapy attempts for inherited metabolic diseases of erythrocytes

Although more than 14 metabolic deficiencies have been identified causing CNSHA, approaches of gene therapy have been done only in a few of them (Table 1). Below, we are including a short description of the different diseases and the attempts addressed.

Among glycolytic defects causing CNSHA, Glucose 6-phosphate dehydrogenase (G6PD) deficiency is the most common genetic disease. More than 400 million people are affected world wide, showing a vast variability of clinical features. G6PD catalyzes the first reaction of the pentose phosphate pathway, in which Glucose 6-phosphate (G6P) is oxidized and Nicotinamide adenine dinucleotide phosphate is reduced (NADPH) resulting in decarboxylation of CO_2 and pentose phosphate. G6PD plays a central role in the cellular physiology as it is the major source of NADPH, required by many essential cellular systems including the antioxidant pathways, nitric oxide synthase, NADPH oxidase, cytochrome p450 system and others. Indeed, G6PD is essential for cell survival. *G6PD* is a 20 kb X-linked gene that maps to the Xq28 region, consisting of 13 exons and 12 introns, which encode a 514 amino acids protein with ubiquitous expression. More than 100 missense mutations in the *G6PD* gene have been identified [14], being most of them single-point mutations causing an amino acid substitution. Frequently, these mutations cause mild symptoms or no disease, except when patients are challenged by increased oxidative stress or fava beans. However, some mutations provoke severe instability of the G6PD and, as a result, lifelong CNSHA with a variable severity [15,16]. Through genetic studies it has been observed that severe clinical manifestations appear preferentially in exons 7, 10 and 11. As *G6PD* is X-linked, the defect is fully expressed in affected males (hemizygotes who inherit the mutation only from the

http://dx.doi.org/10.5772/52531

mother), whereas in homozygous females the mutations are transmitted from both parents. Thereby, female heterozygotes represent a red blood cell mosaic population, causing a wide range clinical picture.

Disease	Gene	Chrom.	Inheritance	Other symptoms	Bone Marrow Transplantation	Gene Therapy
Glucose-6 Phosphate Dehydrogenase Deficiency (G6PD)	*G6PD*	Xq28	X-linked	jaundice, spleno- and hepatomegaly, hemoglobinuria, leukocyte disfunction, and susceptibility to infections		D: C57BL/6 mice P: Transduction of 5-FU treated BM cells with MMLV-hG6PD or MPSV-hG6PD vectors and subsequent transplantation. R: lethally irradiated C57BL/6 mice [5]
Pyruvate Kinase Deficiency (PKD)	*PKLR*	1q21	A.R	Reticulocytosis, splenomegaly, *hidrops foetalis*, and death in neonatal period	D: normal CBA/N$^{+/+}$ mice + 5FU R: CBA Pk-1^{slc}/ PK-1^{slc} mice C: minimal (100 or 400 cGy) [6]	D: WT mice P: Transduction of 5-FU treated BM cells with pMNSM-hLPK retroviral vector and subsequent transplantation R: lethally irradiated mice [7]
					D: normal CBA/N$^{+/+}$ mice R: CBA Pk-1^{slc}/ PK-1^{slc} mice C: no conditioning [8]	D: CBA PK-1^{slc}/PK-1^{slc} mice P: Transgenic rescue using the µLCR-PKLR-hRPK construct [9]
					D: normal Basenji dogs R: PKD Basenji dogs C: sublethal dose (200 cGy) + mycophenolate memofetil + cyclosporine [10]	D: WT mice P: Transduction of Lin^{-}Sca1^{+} BM cells with a MSFV-hRPK retroviral vector and subsequent transplantation R: lethally irradiated WT mice [11]
					D: HLA-identical sister R: PKD severe patient C: busulfan + cyclophosphamide [12]	D: AcB55 mice P: Transduction of Lin^{-}Sca1^{+} BM cells with a MSFV-hRPK retroviral vector and subsequent transplantation R: lethally irradiated AcB55 mice [13]
Glucose Phosphate Isomerase Deficiency (GPI)	*GPI*	19q13.1	A.R	neuromuscular disturbances		
Triose Phosphate Isomerase Deficiency (TPI)	*TPI1*	12p13	A.R	neuromuscular disorders, mental retardation, frecuent infections and death *in utero*		
Hexokinase Deficiency (HK)	*HK1*	10q22	A.R	defects in platelets		
Phosphofructokinase Deficiency (PFK)	*PFKL*	21q22.3	A.R	myopathy, storage disease type VII		
Bisphosphoglycerate Mutase Deficiency (BPGM)	*BPGM*	7q31-q34	A.R	erythrocytosis		
Glutathion Synthetase Deficiency (GSD)	*GSS*	20q11.2	A.R	5-oxoprolinuria, metabolic acidosis, central nervous system impairment		

A.R, autosomic recessive; D, donor; R, receptor; C, conditioning; P, protocol

Table 1. Most Common Erythroid Metabolic Inherited Diseases. BM transplantation and gene therapy approaches

Patients with CNSHA suffer anemia and jaundice, but often tolerate their condition well. However, G6PD variants with low activity are related with alterations in the erythrocyte membrane facilitating its breakdown and causing intravasal hemolysis. These symptoms are often accompanied by spleno- and hepatomegaly and hemoglobinuria. Besides, leukocyte dysfunctions caused by lower concentration of NADPH appear when G6PD activity is below 5% of the normal activity, leading to an immune depression [17]. Vives *et al.* and other groups have also observed an increased susceptibility to infections [18,19].

Preclinical work from Rovira et al demonstrates that *hG6PD* gene transfer into HSCs may be a viable strategy for the treatment of severe G6PD deficiency [5]. Through the transplantation of pluripotent hematopetic stem cells transduced with γ-retroviral vectors carrying the wild type human G6PD cDNA, they achieved a stable and lifelong expression of hG6PD in all the hematopoietic tissues of primary and secondary receptor mice. In this study, transgene expression was driven by the 3' LTR from either the Moloney murine leukemia virus (MMLV) or the myeloproliferative sarcoma virus (MPSV), obtaining an efficient transduction in murine hematopoietic progenitors. The corrected cells were then injected into lethally irradiated syngeneic mice, increasing 2-fold the enzyme activity in peripheral blood cells in comparison with non-transplanted control mice. Long-term hG6PD expression derived from the vector was also observed, which was similar to that of the endogenous enzyme activity. Similar expression was detected in RBC and in White Blood Cells (WBC) in different hematopoietic organs, as expected due to the use of a viral ubiquitous promoter. These results support gene therapy as a suitable strategy for the treatment of severe CNSHA due to G6PD deficiency. Additionally, they also demonstrated the efficacy of this gene therapy vector in human embryonic stem cells (hESC) in which the *G6PD* gene had been inactivated by targeted homologous recombination, which implies the potential application of gene therapy to G6PD hESCs. Moreover, although a selective advantage in favor of G6PD corrected cells has not been reported because the mice used showed normal G6PD activity, Rovira et al observed a strong selection after transduction of G6PD-deficient ES cells with their vectors. In this regard, the development of G6PD deficient mouse models would be a valuable tool to test new protocols. Furthermore, the mouse strain recently developed by Hay Ko et al may be useful, although it does not reproduce all the features of the human G6PD-deficiency [20].

Pyruvate kinase deficiency (PKD), the second most frequent abnormality of glycolysis causing CNSHA, has also been proposed as a potential disease to be treated by gene therapy. Pyruvate kinase (PK) catalyzes the second ATP generation reaction of the glycolysis pathway by converting phosphoenolpyruvate (PEP) into pyruvate, yielding nearly 50% of the total ATP production in red blood cells. PK plays a crucial role in erythrocyte metabolism, since mature RBC are absolutely dependent on the ATP generated by glycolysis, giving the loss of mitochondria, nucleus and endoplasmic reticulum in their mature state. RPK is therefore necessary for maintaining cell integrity and function. Reduced levels of erythrocyte Pyruvate kinase (RPK) lead to an accumulation of glycolytic intermediates that ultimately shortens the life span of mature RBC by metabolic block [21]. Four tissue-specific isoenzymes of PK (M1, M2, R and L) encoded by two different genes (*PK-M* and *PK-LR*) have been identified in humans [22]. The *PK-LR* gene, located on chromosome 1 (1q21) [23] encodes for both LPK (expressed in liver, renal cortex and small

http://dx.doi.org/10.5772/52531

intestine) and RPK (restricted to erythrocytes) through the use of alternative promoters [24]. PK-M1 is expressed in adult nomal tissue, like brain or muscle. The PK-M2 isoform is typically expressed in proliferating tissues like fetal, tumoral and several other adult tissues [25] and during the maturation of the erythroblasts, gradually decreases, giving rise to the RPK isoform.

The codifying region of *PK-LR* gene is split into twelve exons, ten of which are shared by the two isoforms, while exons 1 and 2 are specific for the erythrocyte and the hepatic isoenzyme respectively [26]. However, clinical symptoms caused by *PK-LR* mutations are confined to RBC because the hepatic deficiency is usually compensated by the persistent enzyme synthesis in hepatocytes [27]. To date, more than 150 different mutations in the *PK-LR* have been associated with CNSHA, being most of them missense mutations, splicing and codon stop. Only two variants, -72 G and -83 C, have been identified in the promoter regions so far [26,27]. Molecular studies indicate that severe syndrome is commonly associated with disruptive mutations and missense mutations involving the active site or protein stability [28].

PK deficiency is transmitted as an autosomal recessive trait and although its global incidence is still unknown, it has been estimated in 1:20000 in the general caucasian population [29]. Clinical symptoms appear in homozygous and compound heterozygous patients, which lead to a very variable clinical picture, ranging from mild or fully compensated forms to life-treating neonatal anemia necessitating exchange transfusions and subsequent continuous support [28]. Pathological manifestations are usually observed when enzyme activity falls below 25% of normal PK activity [30], and severe disease has been associated with a high degree of reticulocytosis [31]. *Hydrops foetalis* and death in the neonatal period have also been reported in rare cases [32,33]. PK deficiency treatment is based on supportive measures since no specific therapy for severe cases is available to date. Periodic cell transfusions may be required in severe anemic cases, often impairing their quality of life. Splenectomy can be clinically useful in some patients increasing the hemoglobin levels, as well as iron chelation to decrease the common iron overload observed in PKD patients [34]. However, in some severe cases, allogeneic bone marrow transplantation is required and it has been successfully performed in one severe affected child [12].

The feasibility of gene therapy in PKD was first reported by the group of Asano, who introduced the human LPK cDNA into C57BL/6 mouse bone marrow cells using a retroviral vector [7]. They demonstrated the expression of the LPK transgene mRNA in both peripheral blood and hematopoietic organs after bone marrow transplantation. However, viral-derived expression in peripheral blood was detectable no longer than 30 days post-transplantation, indicating an insufficient transduction efficacy of the retroviral vector used or transduction of non-pluripotent BM cells. In a hemolytic anemia dog model, bone marrow transplantation of minimal conditioned receptors failed to correct the hematological symptoms [10]. Other approaches to rescue RPK phenotype through a gene addition strategy have been also addressed using a PKD transgenic mouse model (*CBA/N PK-1^{SLC}/PK-1^{SLC}*) [9]. In this assay, the hemolytic anemia and reticulocytosis was fully corrected when the human gene was highly expressed by means of pronuclear injection, although splenomegaly was still present. Interestingly, the authors observed a negative correlation between RBC PK activity and the number of apoptotic erythroid progenitors in the spleen, providing evidence that the meta-

bolic alteration in PK deficiency affects not only the survival of RBC, but also the maturation of erythroid progenitors, resulting in ineffective erythropoiesis [35]. Further studies from this group indicate that RPK plays an important role as an antioxidant during erythrocyte differentiation, since glycolytic inhibition by mutations in *Pklr* gene increased the oxidative stress in SLC3 cells (established from *Pk-1*^{slc} mouse) and led to the activation of hypoxia-inducible factor-1 (HIF1), as well as the expression of downstream proapoptotic genes [36].

In addition, our work carried out in mouse models supported the therapeutic potential of viral vectors for the gene therapy of PK deficiency. Throughout the transduction of bone marrow cells using γ-retroviral vectors that carry the human RPK cDNA and subsequent transplantation, we reported a long-term expression of the human protein in RBC obtained from primary and secondary receptor mice, without detectable adverse effects [11]. Recently, we have also reported a successful gene therapy approach using the same retroviral vectors in the congenital mouse strain AcB55, identified by Min-Oo in studies of alleles involved in malaria susceptibility [37]. These mice carry a loss-of-function mutation (269T-> A) resulting in the amino acid substitution I90N in the *Pklr* gene, which yields a similar RBC phenotype to that observed in PKD patients, including splenomegaly and constitutive reticulocytosis. Retroviral-derived expression was capable of fully resolving the pathological phenotype in terms of hematological parameters, anemia, reticulocytosis and splenomegaly, together with normalization of bone marrow and spleen erythroid progenitors, erythropoietin (EPO), PK activity and ATP levels. Interestingly, despite a strong viral promoter was used to drive the expression of the transgene, metabolic energy balance was no modified in white blood cells. Moreover, we observed that values above 25% of genetically corrected cells were needed to fully rescue the deficiency [3], suggesting that RPK transfer protocols will always require a significant extent of gene-complemented HSC. Nevertheless, other experiments performed in the *CBA/N PK-1*^{SLC}/*PK-1*^{SLC} mouse model of PKD have reveled that 10% of normal BM renders RBC expressing nearly normal RPK protein levels [5]. Differences in the genetic defect of the mouse models used could account for these discrepancies, reinforcing the need for high transduction efficiencies to address the disease in the heterogeneous human population. Additionally, we have proposed the *in utero* transplantation of gene corrected cells as an alternative option for the treatment of PKD. The transplantation of RPK deficient lineage negative fetal liver cells transduced with lentiviruses (LVs) expressing the human wild type version of the RPK in 14.5 day-old fetuses partially restored the anemic phenotype, mainly due to a low engraftment of corrected cells [13]. Improved *in utero* cell transfer would allow therapeutic levels, thus offering an alternative therapeutic option for prenatally diagnosed severe PKD. Following our results in the AcB55 mouse model of PKD, phenotype correction could be reached if the percentage of engraftment of corrected cells is significant. We are currently developing improved lentiviral vectors that could be applied in future clinical settings.

Glucose phosphate isomerase (GPI) deficiency is the third most common hereditary cause of CNSHA, due to mutations in *GPI* gene located on the long arm of chromosome 19. The prevalence of this disease is still unknown, with no more than 50 cases reported so far, and with a higher incidence in the black population. The enzyme catalyzes the reversible isomer-

http://dx.doi.org/10.5772/52531

ization from glucose 6-phosphate to fructose 6-phosphate, an equilibrium reaction of the glycolysis pathway. Glucose turnover is affected only in deficiencies below a very low critical residual GPI activity, but with a drastic decline of lactate formation. As no isoenzyme does exist, patients suffer not only from CNSHA and tissue hypoxia, but also from neuromuscular disturbances. In some cases, GPI deficiency has been found in PKD patients, increasing the severity of the clinical scenario and reflecting the degree of the perturbation of glycolysis. The lack of ATP leads to a destabilization of the erythrocyte membrane causing earlier lysis of the RBC and hemolytic anemia of variable degrees [38]. Animal models of GPI deficiency have been described, showing similar symptoms to the human disease [39]. Until now, no gene therapy attempt has been applied to this deficiency.

Other enzyme deficiencies causing CNSHA are Triose phosphate isomerase (TPI) deficiency, associated with neuromuscular disorders, mental retardation and frequent infections, Hexokinase deficiency (HK), affecting also platelet metabolism, phosphofructokinase (PFK) deficiency, 2,3-bisphosphoglycerate mutase (BPGM) deficiency and Glutathione synthetase (GS) deficiency (reviewed in [17,40,41]). Although the incidence of these diseases can be high (ie. TPI is considered as a frequent enzymopathy affecting 0.1% for caucasian populations and even 4.6% for black populations), they are considered rare or very rare diseases, because only few cases (~25 patients in the case of TPI) are diagnosed due to the severity of the clinical manifestations. No gene therapy approaches have been addressed up to now to treat these enzymopaties. However, due to their common characteristics, strategies developed in the other enzyme deficiencies could be applied directly to the treatment of all of them.

3. Optimization of vectors for the gene therapy of metabolic erythroid diseases: Erythroid specific expression vectors

The introduction of a cDNA, encoding for the correct version of the target mutated gene into patient cells using retroviral vectors has been successful for several inherited diseases. The initial integrative vectors for gene therapy design and used in clinical trials were based on Gamma(γ)-retroviral vectors in which the transgene expression was driven by the viral LTR promoter. γ-retroviruses preferably integrate in regions adjacent to the transcription initiation site [42]. The expression of the transgene is promoted by the viral LTR, which drives a high expression that can affect gene regulation of the surrounding genes. Although a high efficiency of transduction and therapeutic effects have been described with these vectors in various monogenic disorders such as immunodeficiencies, adverse effects associated with insertional mutagenesis have also been observed. This has led to the development of the next generation of integrative vectors using self-inactivating-LTR lentiviral backbones. SIN-Lentiviral vectors tend to integrate in intergenic transcribing areas, which represent a safer integrative pattern than γ-retroviral vectors. Aditionally, the expression of the transgene is driven by internal promoters, offering a more physiological expression and a less genotoxic profile when using weak promoters [43]. Current efforts to reduce the mutagenic potential of gene therapy vectors are focussed on not only the use of new viral backbones [44] but also on tissue-specific promoters to restrict the transgene expression to target cells [45] and insu-

lators to confer position-independent expression [46]. Additional regulatory DNA elements such as locus control regions (LCR), enhancers, or silencers have also been used to increase lineage specificity.

Gene therapy for RBC disorders requires, ideally, high erythroid-specific transgene expression in order to avoid side effects in progenitors or hematopoietic lineages other than the erythroid one. In inherited enzymophaties, the overexpression of metabolic enzymes in non-erythroid cells could provide these cells with a potential energetic advantage, with the consequent risk of disturbing the physiological generation of ATP in WBC. Also, the restriction of transcriptional activity to target cells with the use of either tissue-specific or physiologically regulated vectors decreasees the effect of the integrative vectors in the host genome. This goal is particularly important for erythrocyte metabolic deficiencies, as all the affected enzymes are highly regulated and connected with central metabolic pathways. Indeed, an expression limited to the erythroid progeny would reduce the genotoxic risk, as RBC become transcriptionally inactive during differentiation, and finally extrude their nucleus. To study tissue-specific gene therapy strategies for RBC diseases, hemoglobinophaties have been the most widely used.

Erythroid regulatory elements have been extensively used to manage targeted expression to RBC using reporter genes (Table 2). The Locus Control Regions (LCR), defined by their ability to enhance the expression of linked genes to physiological levels in a tissue-specific and copy number-dependent manner at ectopic chromatin sites are commonly used. The components of the LCR normally colocalize to sites of DNase I hypersensitivity (HS) in the chromatin of expressing cells. Individual HS are composed of arrays of multiple ubiquitous and lineage-specific transcription factor-binding sites. In early experiments performed with retroviral backbones, the group of Ferrari developed an erythroid-specific vector by the replacement of the constitutive retroviral enhancer in the U3 region of the 3′ LTR with the HS2 autoregulatory enhancer of the erythroid GATA-1 transcription factor gene. The expression of this vector was restricted to the erythroblastic progeny of both human progenitors and mouse-repopulating stem cells [47,48]. Later, they showed that the addition of the HS1 enhancer to HS2, both from the GATA-1 gene, within the LTR of the retroviral vector significantly improved the expression of the reporter gene. Another enhancer element that has been used to achieve erythroid-specific expression is HS40, located upstream of the ζ-globin gene, since it is able to enhance the activity of heterologous promoters in a tissue-specific manner [49]. It has been shown to be genetically stable in MMLV vectors and enhances expression comparable to that of a single -globin gene [50], although HS40 lacks some of the properties of the LCR, like position independence [51] or copy number dependence [52].

An additional improvement to provide safer vectors for RBC gene therapy was provided by the use of insulators elements, which have been shown to reduce position effects in transgenic animals [60]. Insulators are genomic elements that can shelter genes from their surrounding chromosomal environment, by either blocking the action of a distal enhancer on a promoter [60,61], or by acting as barriers that protect the gene from the silencing effect of heterochromatin [61]. The most well studied element is the chicken hypersensitive site 4

http://dx.doi.org/10.5772/52531

(cHS4), an insulator sequence of the chicken -like globin cluster. Studies performed by Chung et al with the γ-globin promoter and the neo reporter gene on selected cells lines, demonstrated the ability of cHS4 to insulate the expression cassette from the effects of a strong -globin LCR element [63] and therefore reducing its genotoxicity. Experiments from Arumugam et al showed a two-fold reduction in transforming activity with insulated LCR-containing lentiviral vectors comparing with vectors lacking the cHS4 element [68].

Erythroid tissue-specific vectors			
Promoter / enhancer	**transgene**	**Vector type**	**Reference**
HS2 GATA-1 enhancer within the LTR	ΔLNGFR and Neo^R / EGFP	SFCM retroviral vector	[47]
HS1 to HS2 GATA-1 enhancer within the LTR	EGFP and hΔLNGFR	SFCM retroviral vector	[48]
Ankyrin-1 and α-spectrin promoters combined or not with HS40, GATA-1, ARE and intron 8 enhancers	EGFP	HIV-1 based vectors	[53]
α-globin HS40 enhancer and Ankyrin-1 promoter	GFP / *FECH*cDNA	HIV-1 based vectors	[4]
IHK, IHβp and HS3βp chimeric enhancers/ promoters	hβ-globin cDNA	Sleeping beauty transposon	[54]
Physiologically regulated vectors			
Promoter / enhancer	**transgene**	**Vector type**	**Reference**
HSFE and β-globin promoter	hβ-globin cDNA	MSCV retroviral vector	[55]
LCR and β-globin promoter	hβ-globin cDNA or EGFP	HIV-1 based vectors	[56,57]
β-globin and θ-globin promoters combined or not with HS40, GATA-1, ARE and intron 8 enhancers	EGFP	HIV-1 based vectors	[53]
LCR HS4, HS3, HS2, β-globin promoter and truncated β-globin intron 2	EGFP	HIV-1 based vectors	[58]
LCR, cHS4 and β-globin promoter	hβ-globin cDNA	HIV-1 based vectors	[46]
β-globin promoter, LCR HS2, HS3, HS4	hβ-globin cDNA	AAV2	[59]

LTR, long terminal repeats; HS: hypersensitive site; IHK, human *ALAS*2 intron 8 enhancer, HS40 from αLCR and ankyrin-1 promoter; Ihβp, human *ALAS*2 intron 8 enhancer, HS40 from αLCR and β-globin promoter; HS3βp, HS3 core element form human βLCR and β-globin promoter; LCR, locus control region. Modified from Toscano *et al.*, 2011

Table 2. Specific vectors for gene therapy of erythroid inherited diseases.

Tissue-specific expression using alternative human promoters can be convenient or more efficient for some diseases, but driving the expression of the therapeutic genes using own promoters is still the most physiological approach to reduce the genotoxic risk of integrating gene vectors [62]. The use of physiologically regulated vectors has been limited mainly because the promoter and the enhancer elements have to be obtained from the affected genes and they are often too large to be included in a lentiviral backbone, and also because the gene expression pattern depends partially on chromatin positioning [63]. -globin LCR has been widely used when attempting to solve this limitation. The -globin LCR consists of 5 HS regions located upstream of the entire cluster of human -like globin genes, each containing a high density of erythroid-specific and ubiquitous transcription binding elements [64]. Much of the transcriptional activity of the -globin LCR resides in HS2 and HS3 sites, but site 4 is important in adult globin expression [65]. Previous studies *in vitro* and *in vivo* have shown that -globin LCR can enhance erythroid-specific expression from heterologous non-erythroid promoters [66,67]. First approaches using -globin LCR and 3' enhancers were based on murine γretroviral vectors [74,75], but the limited packaging capacity of these vectors (up to 8 kb) did not allow the presence of such as large regulatory sequences. Several vector designs including different combinations of regulatory sequences and a deletion of a cryptic polyadenylation site within intron 2 of -globin gene [68], flanked by an extended promoter sequence and the -globin 3' proximal enhancer were developed. The combination of the LCR elements (3'2 kb) spanning HS2, 3 and 4, were the best amongst several possibilities [69] to achieve a high titer retroviral vector capable of expressing high levels of the transgene.

Other approaches to achieve consistent long-term expression of a transgene have been based on the use of HSFE element, an erythroid-specific chromatin remodelling element derived from the human β-globin LCR which contains binding sites for the erythroid-specific factors NF-E2, GATA-1, EKLF and the ubiquitous factor Sp-1, all of which are necessary to establish a hypersensitive chromatin domain. Work by Nemeth *et al.*, demonstrated that the HSFE can mediate functional tissue-specific "opening" of a minimal human β-globin promoter and increases expression of a human β-globin gene in both MEL cell clones and in transgenic mice. Their results indicated that the most effective vector included tandem copies of the HSFE and produced a 5-fold increase in expression compared to the promoter alone [55] in the context of an integrated retroviral vector.

Gene therapy for RBC metabolic diseases can also benefit from the new technologies based on the modification in mRNA stability or translation efficacy of the transgenes. The use of the post-transcriptional regulatory element (Wpre) from the woodchuck hepatitis virus (WHV) has significantly increased transgene expression in target cells [64,65], even in HSC [70] by stabilization of mRNA at post-transcriptional level. However, it may raise safety concerns, since it contains a truncated form of the WHV *X* gene, which has been implicated in animal liver cancer [71]. Therefore, Wpre has subsequently been improved by a mutation of the open reading frame of the *X* gene [72]. Combination of erythroid promoters like ankyrin-1 or -spectrin with Wpre sequence increased 2-fold the expression in unilineage erythroid cultures [53], and when combined also with erythroid enhancers inserted in tandem: HS40 and GATA-1 or HS40 and I8 enhancers [53]. RNA targeting strategies have mainly been used to down regulate expression of cellular genes using vectors expressing interference RNAs (iRNAs). They can be also used to control the expression of integrating vectors knocking down the transgene by the

http://dx.doi.org/10.5772/52531

endogenous microRNA cellular machinery. Following this strategy, engineered microRNA target sequences in the vector (miRTs-vector) are recognized by a cell specific microRNA (miRNA), avoiding the expression of the therapeutic gene in undesired cell populations [63]. Several miRNAs are differentially expressed during hematopoiesis and their specific expression regulates key functional proteins involved in hematopoietic lineage differentiation. Particularly, miR-223 has been proposed as a myeloid-specific regulator that negatively regulates progenitor proliferation and granulocyte differentiation and activation [73]. Moreover, Felli et al observed that hematopoietic progenitor cells transduced with miR-223 showed a significant reduction of their erythroid clonogenic capacity, suggesting that down-modulation of this miRNA is required for erythroid progenitor recruitment and commitment [79]. Further studies may determine if the use of miRNA-223 target in lentiviral vectors could be useful to achieve a desirable erythroid-specific expression for gene therapy of red blood cell diseases.

In addition, the erythroid-specificity of short segments of the -globin LCR element has been documented in adeno-associated virus 2 (AAV2) system. Their efficacy to mediate an erythroid-restricted expression has been proved by Tan *et al.*, who reported a successful AAV2-mediated high and stable transduction of the human -globin gene in HSCs from -thalassemia mouse model, which were then transplanted into recipient and rescued them of the disease [59]. These vectors have gained attention as potential useful vectors for human gene therapy, mainly because of their non-pathogenic nature in humans and their relativly easy production. Besides, AAV2 vectors are easily purified to high titers and are able to transduce dividing and non-dividing cells. However, most of proviral AAV2 genomes remain episomal and the insert size is restricted to just over 4kb. Further studies are still needed to know whether they would be a better option than current lentiviral vectors. Also, long-term genotoxic risk of recombinant AAV2 therapy in human is not known up to the date.

In addition, the efficacy of some of these erythroid-specific elements and promoters has also been tested in non-viral vectors, such as transposons. Zhu et al, for instance, studied several hybrid promoters driving the expression of the human -globin gene using the sleeping beauty transposon (SB-Tn). They combined several erythroid elements to develop different chimeric promoters. Their results indicated that the ankyrin-1 minimum promoter was stronger than -globin's, and the hALAS I8 enhancer (IH) was significantly more powerful that HS3 core element from -LCR and -globin promoter [54]. SB-Tn system is a promising non-viral vector for efficient genomic insertion, even with erythroid-specificity. However, its efficiency for delivering transgenes into HSCs is still much lower than other engineering viral vectors.

4. Overcoming conventional gene therapy pitfalls: gene editing in induced plutipotent stem cells

4.1. Human induced pluripotent stem cells and reprogramming platforms

Since Yamanaka et al first reported the generation of mouse induced Pluripotent Stem Cells (iPSC) in 2006 by the ectopic expression of four transcription factors (Oct4, Sox2, Klf4 and

cMyc) [74] and one year later in human cells together with Thompson's group [75,76], many laboratories around the world have been able to reprogram a large range of somatic cells into pluripotent stem cells, from neural stem cells [77] to terminally differentiated B-lymphocytes [78]. The reproducibility and potentiallity (unlimited self-renewal and ability to differentiate into any cell type) of this technology has made the iPSC field to advance very rapidly. The human iPSC (hiPSC) technology brings together all the potential of hESC in terms of pluripotency without any ethical issue and the immunotolerance of the autologous cell treatment. Therefore, hiPSC technology is one of the most promising fields for future therapies for many human diseases. Safer reprogramming approaches have been designed and many patient specific hiPSC have been generated both to model human diseases and to correct by gene therapy approaches. Depending on the cell type to be reprogrammed, the number of factors used could be reduced and, what is more important, oncogenes or tumor related proteins included in the reprogramming cocktail, like *c-MYC* or *KLF4* [79] could be removed from the original reprogramming cocktail [80-82]. Several groups developed excisable polycistronic lentiviral vectors [83,84] or transposon-based reprogramming systems [85,86], which could be removed after getting the hiPSC clones. Similar results have been obtained using recombinant proteins [82], synthetic mRNAs [87], and non integrating RNA Sendai Virus vectors [88]. Except for Sendai viruses, non integrating methods show a reduced reprogramming efficiency and the range of cells reprogrammed is not as large as with lentiviral or retroviral vectors.

iPSC technology makes feasible the availability of patient specific cells to study the biology of the disease and develop advanced tools to cure the phenotype and could potentially be used as a therapeutic option (Figure 1). Focussing on metabolic diseases, the first reported metabolic disease patient specific hiPSC line was obtained one year after the first generation of hiPSCs. It was from a 42-year old female that suffered from Type I Diabetes mellitus [89] and it showed no differences compared to a wild type hiPSC line in terms of pluripotency. Next report in which liver metabolic disease patient samples were reprogrammed was carried out by the group of Ludovic Vallier [90], and showed the potential of this kind of approaches for disease modelling and new drug discovery. They reprogrammed fibroblast obtained from α-1 Antitrypsin deficiency (*A1ATD*), Familiar Hypercholesterolemia (FH), Glucose-6-Phosphate deficiency (G6PD), Crigler-Najjar Syndrome and hereditary Tyrosinemia Type 1 patients, and generated hepatocytes that showed characteristics of mature hepatic cells, like albumin secretion or cytocrome p450 metabolism. Three of the five cell lines (*A1ATD*, FH, and GSD1a hiPSCs) were capable of recapitulating the disease phenotype in vitro. Disease modelling in erythroid diseased induced pluripotent cell lines has been performed for -Thalassemia [91,92] and sickle cell anemia [93,94]. In these reports the phenotype was corrected by LVs integrated in areas of the genome that were considered safe for viral integration [83] or by gene editing using homologous recombination of the affected locus [91,93,94].

The future therapeutic application of hiPSC will not only require non-integrative reprogramming system, but also a more precise gene correction. During last years, the cooperation between hiPSC technology and gene editing is being explored. Human iPSC technology has

http://dx.doi.org/10.5772/52531

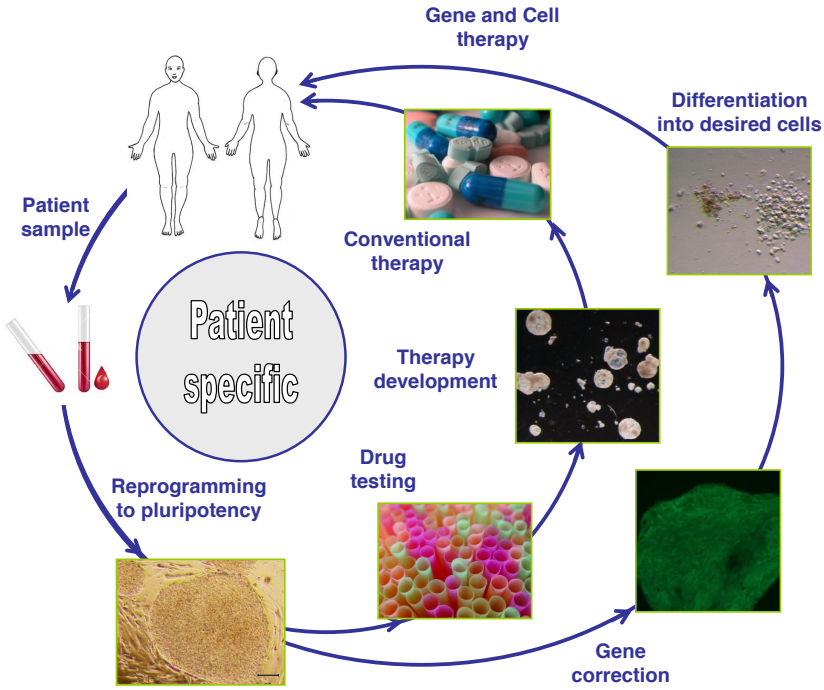

Figure 1. Potential utilities of hiPSC and iPS technology

led to the opportunity to control the integration of viral vectors at a clonal level. As we have mentioned before, the analysis of lentiviral integration sites in β-thalassemia hiPSC allowed the identification of corrected hiPSC clones expressing β-globin transgene from a safe genomic site (also called Safe harbour), a site in which integration does not disturb the expression of any neighbouring genes during their erythroid differentiation [83]. The therapeutic use of patient-specific hiPSC emerges then from the combination of gene and cell therapy. From this new research field,future gene therapy protocols will emerge.

4.2. Gene editing based on homologous recombination

Gene editing is a process in which a DNA sequence is introduced into a specific locus or a chromosomal sequence is replaced. This site-specific precise introduction requires an accurate recognition mechanism of the target site on the genome. Under normal conditions, the maintenance of the integrity of the genome requires that the cells repair DNA damage with high fidelity. One of the most harmful DNA damage is the generation of double-strand breaks (DSB). DSB are often resolved by non-homologous end joining (NHEJ), which joins the two ends of the DSB. However this DNA repair mechanism could introduce mutations. On the contrary, homologous recombination (HR) is a truly accurate DNA repair mechanism because it is basically a “copy and paste” mechanism. This process uses an undamaged

homologous segment of DNA that can be exogenously provided as a template to copy the information across the DSB. The fidelity of HR gives us the specificity and accuracy that gene editing requires.

The natural HR process has been adapted by researchers to get the desirable addition of an exogenous cassette into the targeted locus. This techniques have been widely use for the generation of knock-out and knock-in transgenic animals [95]. To correct or insert and express a transgene by HR we can consider three different strategies: i) Gene correction, a base or some bases can be substituted from the original strand using an homologous sequence where this base or bases are modified; it is the way to introduce/repair point or small mutations; ii) Safe harbour integration, a complete expression cassette (promoter, transgene and regulatory signals) is inserted in a safe place of the genome, without altering the expression of the surrounding genes and without being silenced by epigenetic mechanisms; this is the case for *AASV1* and *CCR5* loci. Additionally to these well known safe harbours, there is a wide research focused on finding potential new safe harbour places. iii) Knock-in insertion, the cDNA of a gene is introduced in the same site of the endogenous gene, linked by splicing mechanisms to the endogenous gene assuring the expression of the inserted sequence by the endogenous regulatory elements of the locus where it is integrated.

Gene editing process can be separated in two different steps, generation of DSB and HR. The efficacy of gene editing in human cells depends on the generation of DSB at the specific target site and on the DNA repair mechanism that the cell uses to resolve the DSB. Unfortunately, NHEJ is the dominant pathway to solve these DNA lesions in human cells. Additionally, HR varies in different cell types and requires transit through S-G2 phase of the cell cycle [96]. These limitations make gene editing in human cells difficult to achieve. However, different approaches are being used to improve gene editing by HR, like increasing the length of the DNA sequences homologous to target site (homology arms) [97], the use of adeno-associated vectors [98], the improvement of selection methods for edited cells or the stimulation of HR by inducing DSB using DNA nucleases.

Recently, engineered DNA nucleases have been developed to specifically induce DSB at a unique and defined sequence in the cell genome. These proteins are formed by a nuclease domain and a DNA binding domain whose sequence specificity can be engineered. The most widely used DNA nucleases are Zinc finger nucleases (ZFN), homing meganucleases (MN) and transcription activator-like effector nucleases (TALEN). They identify a potentially unique sequence in the genome and generate DSBs in the desired genomic site, aiming to promote the repair of the DSB by the cell machinery and, ideally by HR. The DNA binding domain of a ZFNs is derived from zinc-finger proteins and is linked to the nuclease domain of the restriction enzyme Fok-I. DNA-biding domain is a tandem repeat of Cys_2His_2 zinc fingers, each of which recognizes three nucleotides. ZFNs work as pairs of two monomers of ZFN, one in reverse orientation. This ZFN dimer can be designed to bind to genomic sequences of 18-36 nucleotides long. TALENs have a similar structure to ZFNs, but the DNA-binding domain comes from transcription activator-like effector proteins. The DNA-binding domain in TALENs is a tandem array of amino acid repeats. Each of these units is able to bind to one of the four possible nucleotides and this makes that the DNA binding domain can be designed to recognize any desired genomic sequence. TALENs also cleave as dimers. Contrary to these synthetic DNA-nucleases, MNs are a

http://dx.doi.org/10.5772/52531

subset of homing endonucleases which recognize a DNA sequence from 14 to 40 nucleotides. Current MNs have been engineered from natural homing endonucleases to increase the number of target DNA sequences.

ZNFs have been widely used for gene editing in hESC and hiPSC. In 2007, Dr. Naldini's laboratory showed the insertion of GFP into the CCR5 safe harbour in human stem cells (HSC and hESC) after inducing HR by ZFN expression. The CCR5-ZFN and donor DNAs were delivered into hESC by intergrative deficient lentiviruses. More interestingly, targeted hESC were able to differentiate into neurons keeping GFP expression [99]. Soon, the proof of principle for the clinical application of ZFN-mediated gene editing was tested in hiPSC from patients affected by different genetic diseases. The first pre-clinical use of ZFN for gene therapy of a metabolic disease was performed by Yusa *et al*. In this report, gene correction was performed at the α1-antitrypsin (*A1AT*) locus to revert A1AT deficiency in hiPSC derived from a patient with a point mutation. This group included a Puromycin resistence cassette flanked by piggyBac sites, so that the Puromycin selection facilitated the isolation of corrected A1ATD-iPSC clones. Afterwards, the selection cassette was removed by piggyBac transposon, obtaining corrected hiPS clones without any additional sequence. These corrected hiPS clones were then differentiated into hepatocyte-like cells to confirm the complete correction of the A1ATD [101]. Other hiPSC gene editing approaches and functional correction of erithroid diseases include gene correction of Sickle Cell Anemia [94] and -Thalassemia [91].

One of the major limitations of ZFN is the generation of "off-target" DSB, due to unspecific sequence recognition. Different studies have highlighted this as a possible limitation in the clinical use of ZFN-mediated HR [100,101]. Recent works have explored the potential of other types of DNA-nucleases in order to prevent the "off-target" cleave limitations of the ZFN, being TALEN and MN the most promising ones. The feasibility of TALEN to mediate HR in hESC and hiPSC was assess by Jaenisch's group when they designed TALEN targeting the *PPP1R12C* (at *AAVS1* locus), *POU5F1* and *PITX3* genes at precisely the same positions as the one targeted by ZFN in their previous work [102]. The authors described a gene editing efficiency similar to the one achieved by ZFN with a low level of "off-targets" [103]. A strategy to minimize the potential number of "off-targets" is to design TALEN to work as obligatory heterodimers, which has beeing already done in the engineered MNs. The application of the TALEN and MN as tools to improve HR is still on going. We are exploring the pre-clinical use of TALEN and MN to correct erythroid metabolic genetic diseases, such as PKD.

5. Complementary developments for the application of gene therapy to erythroid metabolic diseases

5.1. *In vivo* transduction using engineered envelopes

Another challenge for the clinical application of gene therapy relates to vector targeting. To achieve successful gene therapy, the appropriate gene must be delivered to target cells and specifically expressed in them, without harming non-targeted cells. The most common and easiest way to target specific cells is by *ex vivo* infection of the desired cell population. There-

fore, cells can be directly exposed to the viral vectors facilitating viral-cell interaction. These interactions are driven by the envelope protein which can be adapted from other viruses redirecting the tropism of the vector. The most widely used vectors are lentiviral vectors pseudotyped with the attachment glycoprotein of the vesicular stomatitis virus (VSV-G), which allows the production of high-titre vectors and confers a broad host range [104]. In comparison with them, engineered LVs capable of delivering genes of interest to predetermined cells, can reduce the targeting of undesirable cell types and improve the safety profile, which will further enhance the use of this vector system for gene therapy applications [105,106]. As we have mentioned above, the use of promoters and regulatory sequences that are only active in target cells adds lineage specific expression, although integration of the viruses in non desired cells is still possible. *Ex vivo*-targeted gene delivery, as commonly used in HSCs transduction, is associated with a risk of inducing cell differentiation and loss of the engraftment potential of these cells [107]. On the contrary, *in vivo* gene transfer could target HSCs in their stem cell niche, a microenvironment that regulates HSC survival and maintenance [105]. To accomplish this, the vector must display a suitable system to selectively infect the desired population, for example the introduction of a specific ligand to bind a target-cell receptor [106].

Many attempts have been made to develop targeted transduction systems using retroviral and lentiviral vectors by altering the envelope glycoprotein (Env), which is responsible for the binding of the virus to the cell surface receptors and for mediating viral entry into the cell. The plasticity of the surface domain of Env allows insertion of ligands, peptides or single-chain antibodies that can direct the vectors to specific cell types [108]. However, this type of manipulation negatively affects the fusion domain of Env, resulting in low viral titers. To overcome this downside, a method to engineer lentiviral vectors has been developed. These vectors transduce specific cell types by breaking up the binding and fusion functions of the envelope protein into two distinct proteins [108]. Instead of pseudotyping lentiviral vectors with a modified viral envelope protein, these lentiviral vectors co-display a targeting antibody and a fusogenic molecule on the same viral vector surface. Based on molecular recognition, the targeting antibody should direct lentiviral vectors to the specific cell type. The binding between the antibody and the corresponding cellular antigen should induce endocytosis resulting in the transport of lentiviral vectors into the endosomal compartment. Once inside the endosome, the fusogenic molecule should undergo a conformational change in response to the decrease in pH, thereby releasing the viral core into the cytosol [109]. The use of fusion domain of the binding defective Sindbis virus glycoprotein together with an anti-CD20 antibody has been shown to mediate the targeted transduction of lentiviral vectors to CD20-expressing B cells [110].

However, two major challenges for *in vivo* gene delivery are LVthe exposure to the host immune/complement system and off-target cell transfer after systemic administration. For these reasons, second generation of early-acting-cytokine-displaying LVs has been developed, that circumvents these obstacles by specifically targeting $hCD34^+$ cells [111,112]. For example, RDTR/SCFHA-LV, consisting of RD114 glycoprotein and stem cell factor (SCF) fused to the *Influenza hameglutinin* env protein, is resistant to degradation by human comple-

http://dx.doi.org/10.5772/52531

ment and efficiently transduces very immature hCD34$^+$ HSCs [113]. This new generation of HSC-targeted LVs should improve current gene therapy protocols through the transduction of primitive HSCs directly in the bone marrow of patients with genetic diseases.

5.2. *In vitro* production of mature erythrocytes

Periodical blood transfusion is the previous to the last therapeutic option for severe cases of CNSHA patients. However, this clinical practice involves also adverse effects related to the immuneresponse against minor erythrocyte antigens which makes the patients refractory to additional blood transfusions in the long run. The availability of genetically corrected patient-specific iPSC would allow the possibility of generating disease free erythrocytes ready for transfusion, avoiding the adverse immune effects.

There have been numerous attempts to produce RBC *in vitro* from different sources of stem cells. To date, the most successful protocol has been developed by the group of Luc Douay [113,114]. Using peripheral blood CD34$^+$ cells, these authors were able to expand and generate RBC with *in vitro* and *in vivo* features of native RBC, and were also capable of transfusing a patient with *in vitro* generated erythrocytes. Notably, the same authors reported a protocol to generate RBC from hiPSC as an alternative source of HSC [114]. Other groups have described similar protocols to generate erythrocytes from hESC or hiPSC [115-118], although in all these studies the RBC generated from embryonic cells expressed embryonic and foetal hemoglobins but low levels of adult hemoglobin. Additional efforts should be done to make this possibility a therapeutic option.

6. Conclusions

Erythroid metabolic diseases are well defined and well known diseases which main symptom is CNSHA. As they are monogenic diseases that can be cured by allogeneic bone marrow transplantation, they are very good candidates to be treated by gene therapy. However, the low number of patients with poor prognosis requiring BM transplantation and the absence of an apparent selective advantage of the corrected cells over the diseased ones have made their approach for gene therapy less attractive than other erythropaties. Up to now, no gene therapy clinical trial for erythroid metabolic diseseases has been accomplished. Gene therapy attempts in animal models have been applied to G6PD and PKD with successful results, emphasizing the usefulness of a gene therapy approach for these diseases. Although adverse effects due to ectopic expression of the metabolic enzyme have not been observed, an erythroid specific expression is preferred. Many developments have been made for the specific expression of globin genes that could be adapted to vectors developed for the discussed erythroid metabolic diseases. Similarly, any attempt directed to the improvement of HSC transduction, including the possibility of *in vivo* targeted gene therapy could be applied. On the other hand, the combination of cell reprogramming and gene editing opens a new world of possibilities that could be easily applied to these diseases. hESC and hiPSC are helping in the development of the next generation of gene therapy, which implies a precise

gene targeting. Gene editing by HR is the best and safest gene therapy procedure because avoids any perturbation in the targeted genome. Besides the combination of hiPSC and gene editing could be the future therapy for many genetic-based diseases. The hiPSC technology is the springboard for the development of more efficient HR protocols applicable to other types of stem cells such as hematopoietic stem cells. The combination of methods for obtaining big amounts of RBC from HSC or embryonic cells, along with the improvement of the different gene therapy approaches described in this chapter, opens up the possibility of the therapeutic application involving the infusion of RBC differentiated *in vitro* from genetically corrected patient specific stem cells.

Nomenclature

5-FU 5-fluorouracil

A1ATD-1 antitripsin deficiency

AAV Adeno-associated virus

BM Bone marrow

BPGM 2,3-bisphosphoglycerate mutase

CNSHA Chronic non spherocytic hemaolotyc anemia

DSB Double strand breaks

Env Viral envelope

FH familiar hypercholesterolemia

G6P Glucose-6-phosphate

G6PD Glucose-6-phopahate dehydrogenase

GPI Glucose phosphate isomerase

GS Glutathione synthetase

hESC human embryonic stem cell

hIF1 hypoxia-inducible factor-1

hiPSC Human induced pluripotent stem cell

HK Hexokinase

HR Homologous recombination

HS DNase I hypersensitive sites

HSC Hematopoietic stem cell

iPSC Induced pluripotent stem cell

http://dx.doi.org/10.5772/52531

kb kilobases

LCR Locus control region

LTR Long terminal repeats

LV Lentivirus

MN homing meganuclease

NHEJ non-homologous end joining

PFK phosphofructokinase

RBC Erythrocytes

SIN-LV Self-inactivated lentiviral vector

TALEN transcription activator-like effector nuclease

TPI Triose phosphate isomerase

WT wild-type

ZFN zinc finger nuclease

Aknowledgements

The authors thank L. Cerrato, M.A. Martín and I. Orman for their technical assistance. We would also like to thank Dr. J. Bueren for careful reading and suggestion of the manuscript. M.G.G. was partially supported by a short-term fellowship from the European Molecular Biology Organization (EMBO ASTF 188.00-2010). Z.G. is a fellowship of the PhD program of the Departamento de Educación, Universidades e Investigación del Gobierno Vasco. This work was funded by grants from the Ministerio de Economía y Competitividad (SAF2011-30526-C02-01), Fondo de Investigaciones Sanitarias (RD06/0010/0015) and the PERSIST European project. The authors also thank the Fundación Botín for promoting translational research at the Hematopoiesis and Gene Therapy Division-CIEMAT/CIBERER.

Author details

Maria Garcia-Gomez, Oscar Quintana-Bustamante, Maria Garcia-Bravo, S. Navarro, Zita Garate and Jose C. Segovia

Differentiation and Cytometry Unit, Hematopoiesis and Gene Therapy Division, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) and Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBER-ER), Madrid, Spain

References

[1] Aiuti A, Bachoud-Levi AC, Blesch A, *et al.* Progress and prospects: gene therapy clinical trials (part 2). Gene Ther. 2007;14:1555-1563.

[2] Herzog RW, Cao O, Srivastava A. Two decades of clinical gene therapy success is finally mounting. Discov Med. 2010;9:105-111.

[3] Naldini L. Ex vivo gene transfer and correction for cell-based therapies. Nat Rev Genet. 2011;12:301-315.

[4] Richard E, Mendez M, Mazurier F, *et al.* Gene therapy of a mouse model of protoporphyria with a self-inactivating erythroid-specific lentiviral vector without preselection. Mol Ther. 2001;4:331-338.

[5] Rovira A, De Angioletti M, Camacho-Vanegas O, *et al.* Stable in vivo expression of glucose-6-phosphate dehydrogenase (G6PD) and rescue of G6PD deficiency in stem cells by gene transfer. Blood. 2000;96:4111-4117.

[6] Richard RE, Weinreich M, Chang KH, Ieremia J, Stevenson MM, Blau CA. Modulating erythrocyte chimerism in a mouse model of pyruvate kinase deficiency. Blood. 2004;103:4432-4439.

[7] Tani K, Yoshikubo T, Ikebuchi K, *et al.* Retrovirus-mediated gene transfer of human pyruvate kinase (PK) cDNA into murine hematopoietic cells: implications for gene therapy of human PK deficiency. Blood. 1994;83:2305-2310.

[8] Morimoto M, Kanno H, Asai H, *et al.* Pyruvate kinase deficiency of mice associated with nonspherocytic hemolytic anemia and cure of the anemia by marrow transplantation without host irradiation. Blood. 1995;86:4323-4330.

[9] Kanno H, Utsugisawa T, Aizawa S, *et al.* Transgenic rescue of hemolytic anemia due to red blood cell pyruvate kinase deficiency. Haematologica. 2007;92:731-737.

[10] Zaucha JA, Yu C, Lothrop CD, Jr., *et al.* Severe canine hereditary hemolytic anemia treated by nonmyeloablative marrow transplantation. Biol Blood Marrow Transplant. 2001;7:14-24.

[11] Meza NW, Quintana-Bustamante O, Puyet A, *et al.* In vitro and in vivo expression of human erythrocyte pyruvate kinase in erythroid cells: a gene therapy approach. Hum Gene Ther. 2007;18:502-514.

[12] Tanphaichitr VS, Suvatte V, Issaragrisil S, *et al.* Successful bone marrow transplantation in a child with red blood cell pyruvate kinase deficiency. Bone Marrow Transplant. 2000;26:689-690.

[13] Meza NW, Alonso-Ferrero ME, Navarro S, *et al.* Rescue of pyruvate kinase deficiency in mice by gene therapy using the human isoenzyme. Mol Ther. 2009;17:2000-2009.

[14] Bulliamy T, Luzzatto L, Hirono A, Beutler E. Hematologically important mutations: glucose-6-phosphate dehydrogenase. Blood Cells Mol Dis. 1997;23:302-313.

[15] Beutler E, Kuhl W, Gelbart T, Forman L. DNA sequence abnormalities of human glucose-6-phosphate dehydrogenase variants. J Biol Chem. 1991;266:4145-4150.

[16] Mason PJ, Sonati MF, MacDonald D, *et al.* New glucose-6-phosphate dehydrogenase mutations associated with chronic anemia. Blood. 1995;85:1377-1380.

[17] Jacobasch G, Rapoport SM. Hemolytic anemias due to erythrocyte enzyme deficiencies. Mol Aspects Med. 1996;17:143-170.

[18] Vives Corrons JL, Feliu E, Pujades MA, *et al.* Severe-glucose-6-phosphate dehydrogenase (G6PD) deficiency associated with chronic hemolytic anemia, granulocyte dysfunction, and increased susceptibility to infections: description of a new molecular variant (G6PD Barcelona). Blood. 1982;59:428-434.

[19] Roos D, van Zwieten R, Wijnen JT, *et al.* Molecular basis and enzymatic properties of glucose 6-phosphate dehydrogenase volendam, leading to chronic nonspherocytic anemia, granulocyte dysfunction, and increased susceptibility to infections. Blood. 1999;94:2955-2962.

[20] Ko CH, Li K, Li CL, *et al.* Development of a novel mouse model of severe glucose-6-phosphate dehydrogenase (G6PD)-deficiency for in vitro and in vivo assessment of hemolytic toxicity to red blood cells. Blood Cells Mol Dis. 2011;47:176-181.

[21] Zanella A, Fermo E, Bianchi P, Chiarelli LR, Valentini G. Pyruvate kinase deficiency: the genotype-phenotype association. Blood Rev. 2007;21:217-231.

[22] Fothergill-Gilmore LA, Michels PA. Evolution of glycolysis. Prog Biophys Mol Biol. 1993;59:105-235.

[23] Satoh H, Tani K, Yoshida MC, Sasaki M, Miwa S, Fujii H. The human liver-type pyruvate kinase (PKL) gene is on chromosome 1 at band q21. Cytogenet Cell Genet. 1988;47:132-133.

[24] Noguchi T, Yamada K, Inoue H, Matsuda T, Tanaka T. The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters. J Biol Chem. 1987;262:14366-14371.

[25] Guguen-Guillouzo C, Szajnert MF, Marie J, Delain D, Schapira F. Differentiation in vivo and in vitro of pyruvate kinase isozymes in rat muscle. Biochimie. 1977;59:65-71.

[26] Kanno H, Fujii H, Miwa S. Structural analysis of human pyruvate kinase L-gene and identification of the promoter activity in erythroid cells. Biochem Biophys Res Commun. 1992;188:516-523.

[27] Nakashima K, Miwa S, Fujii H, *et al.* Characterization of pyruvate kinase from the liver of a patient with aberrant erythrocyte pyruvate kinase, PK Nagasaki. J Lab Clin Med. 1977;90:1012-1020.

[28] Zanella A, Fermo E, Bianchi P, Valentini G. Red cell pyruvate kinase deficiency: molecular and clinical aspects. Br J Haematol. 2005;130:11-25.

[29] Beutler E, Gelbart T. Estimating the prevalence of pyruvate kinase deficiency from the gene frequency in the general white population. Blood. 2000;95:3585-3588.

[30] Diez A, Gilsanz F, Martinez J, Perez-Benavente S, Meza NW, Bautista JM. Life-threatening nonspherocytic hemolytic anemia in a patient with a null mutation in the PKLR gene and no compensatory PKM gene expression. Blood. 2005;106:1851-1856.

[31] Miwa S, Kanno H, Fujii H. Concise review: pyruvate kinase deficiency: historical perspective and recent progress of molecular genetics. Am J Hematol. 1993;42:31-35.

[32] Gilsanz F, Vega MA, Gomez-Castillo E, Ruiz-Balda JA, Omenaca F. Fetal anaemia due to pyruvate kinase deficiency. Arch Dis Child. 1993;69:523-524.

[33] Ferreira P, Morais L, Costa R, *et al*. Hydrops fetalis associated with erythrocyte pyruvate kinase deficiency. Eur J Pediatr. 2000;159:481-482.

[34] Zanella A, Bianchi P, Iurlo A, *et al.* Iron status and HFE genotype in erythrocyte pyruvate kinase deficiency: study of Italian cases. Blood Cells Mol Dis. 2001;27:653-661.

[35] Aizawa S, Kohdera U, Hiramoto M, *et al.* Ineffective erythropoiesis in the spleen of a patient with pyruvate kinase deficiency. Am J Hematol. 2003;74:68-72.

[36] Aisaki K, Aizawa S, Fujii H, Kanno J, Kanno H. Glycolytic inhibition by mutation of pyruvate kinase gene increases oxidative stress and causes apoptosis of a pyruvate kinase deficient cell line. Exp Hematol. 2007;35:1190-1200.

[37] Min-Oo G, Fortin A, Tam MF, Nantel A, Stevenson MM, Gros P. Pyruvate kinase deficiency in mice protects against malaria. Nat Genet. 2003;35:357-362.

[38] Lakomek M, Winkler H. Erythrocyte pyruvate kinase- and glucose phosphate isomerase deficiency: perturbation of glycolysis by structural defects and functional alterations of defective enzymes and its relation to the clinical severity of chronic hemolytic anemia. Biophys Chem. 1997;66:269-284.

[39] Merkle S, Pretsch W. Glucose-6-phosphate isomerase deficiency associated with nonspherocytic hemolytic anemia in the mouse: an animal model for the human disease. Blood. 1993;81:206-213.

[40] Hoyer JD, Allen SL, Beutler E, Kubik K, West C, Fairbanks VF. Erythrocytosis due to bisphosphoglycerate mutase deficiency with concurrent glucose-6-phosphate dehydrogenase (G-6-PD) deficiency. Am J Hematol. 2004;75:205-208.

[41] Njalsson R. Glutathione synthetase deficiency. Cell Mol Life Sci. 2005;62:1938-1945.

[42] Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science. 2003;300:1749-1751.

[43] Modlich U, Navarro S, Zychlinski D, *et al.* Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. Mol Ther. 2009;17:1919-1928.

[44] Montini E, Cesana D, Schmidt M, et al. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat Biotechnol. 2006;24:687-696.

[45] Montini E, Cesana D, Schmidt M, *et al.* The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. J Clin Invest. 2009;119:964-975.

[46] Puthenveetil G, Scholes J, Carbonell D, et al. Successful correction of the human beta-thalassemia major phenotype using a lentiviral vector. Blood. 2004;104:3445-3453.

[47] Grande A, Piovani B, Aiuti A, Ottolenghi S, Mavilio F, Ferrari G. Transcriptional targeting of retroviral vectors to the erythroblastic progeny of transduced hematopoietic stem cells. Blood. 1999;93:3276-3285.

[48] Testa A, Lotti F, Cairns L, *et al.* Deletion of a negatively acting sequence in a chimeric GATA-1 enhancer-long terminal repeat greatly increases retrovirally mediated erythroid expression. J Biol Chem. 2004;279:10523-10531.

[49] Ren S, Wong BY, Li J, Luo XN, Wong PM, Atweh GF. Production of genetically stable high-titer retroviral vectors that carry a human gamma-globin gene under the control of the alpha-globin locus control region. Blood. 1996;87:2518-2524.

[50] Emery DW, Chen H, Li Q, Stamatoyannopoulos G. Development of a condensed locus control region cassette and testing in retrovirus vectors for A gamma-globin. Blood Cells Mol Dis. 1998;24:322-339.

[51] Robertson G, Garrick D, Wu W, Kearns M, Martin D, Whitelaw E. Position-dependent variegation of globin transgene expression in mice. Proc Natl Acad Sci U S A. 1995;92:5371-5375.

[52] Sharpe JA, Summerhill RJ, Vyas P, Gourdon G, Higgs DR, Wood WG. Role of upstream DNase I hypersensitive sites in the regulation of human alpha globin gene expression. Blood. 1993;82:1666-1671.

[53] Moreau-Gaudry F, Xia P, Jiang G, *et al*. High-level erythroid-specific gene expression in primary human and murine hematopoietic cells with self-inactivating lentiviral vectors. Blood. 2001;98:2664-2672.

[54] Zhu J, Kren BT, Park CW, Bilgim R, Wong PY, Steer CJ. Erythroid-specific expression of beta-globin by the sleeping beauty transposon for Sickle cell disease. Biochemistry. 2007;46:6844-6858.

[55] Nemeth MJ, Lowrey CH. An Erythroid-Specific Chromatin Opening Element Increases beta-Globin Gene Expression from Integrated Retroviral Gene Transfer Vectors. Gene Ther Mol Biol. 2004;8:475-486.

[56] May C, Rivella S, Callegari J, *et al.* Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. Nature. 2000;406:82-86.

[57] Pawliuk R, Westerman KA, Fabry ME, et al. Correction of sickle cell disease in transgenic mouse models by gene therapy. Science. 2001;294:2368-2371.

[58] Hanawa H, Yamamoto M, Zhao H, Shimada T, Persons DA. Optimized lentiviral vector design improves titer and transgene expression of vectors containing the chicken beta-globin locus HS4 insulator element. Mol Ther. 2009;17:667-674.

[59] Tan M, Qing K, Zhou S, Yoder MC, Srivastava A. Adeno-associated virus 2-mediated transduction and erythroid lineage-restricted long-term expression of the human beta-globin gene in hematopoietic cells from homozygous beta-thalassemic mice. Mol Ther. 2001;3:940-946.

[60] Lisowski L, Sadelain M. Current status of globin gene therapy for the treatment of beta-thalassaemia. Br J Haematol. 2008;141:335-345.

[61] Sun FL, Elgin SC. Putting boundaries on silence. Cell. 1999;99:459-462.

[62] Zychlinski D, Schambach A, Modlich U, *et al.* Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol Ther. 2008;16:718-725.

[63] Toscano MG, Romero Z, Munoz P, Cobo M, Benabdellah K, Martin F. Physiological and tissue-specific vectors for treatment of inherited diseases. Gene Ther. 2011;18:117-127.

[64] Levings PP, Bungert J. The human beta-globin locus control region. Eur J Biochem. 2002;269:1589-1599.

[65] Navas PA, Peterson KR, Li Q, McArthur M, Stamatoyannopoulos G. The 5'HS4 core element of the human beta-globin locus control region is required for high-level globin gene expression in definitive but not in primitive erythropoiesis. J Mol Biol. 2001;312:17-26.

[66] Blom van Assendelft G, Hanscombe O, Grosveld F, Greaves DR. The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. Cell. 1989;56:969-977.

[67] Montiel-Equihua CA, Zhang L, Knight S, *et al.* The beta-Globin Locus Control Region in Combination With the EF1alpha Short Promoter Allows Enhanced Lentiviral Vector-mediated Erythroid Gene Expression With Conserved Multilineage Activity. Mol Ther. 2012;20:1400-1409.

[68] Sadelain M, Wang CH, Antoniou M, Grosveld F, Mulligan RC. Generation of a high-titer retroviral vector capable of expressing high levels of the human beta-globin gene. Proc Natl Acad Sci U S A. 1995;92:6728-6732.

[69] Sadelain M, Rivella S, Lisowski L, Samakoglu S, Riviere I. Globin gene transfer for treatment of the beta-thalassemias and sickle cell disease. Best Pract Res Clin Haematol. 2004;17:517-534.

[70] Ramezani A, Hawley TS, Hawley RG. Lentiviral vectors for enhanced gene expression in human hematopoietic cells. Mol Ther. 2000;2:458-469.

[71] Kingsman SM, Mitrophanous K, Olsen JC. Potential oncogene activity of the woodchuck hepatitis post-transcriptional regulatory element (WPRE). Gene Ther. 2005;12:3-4.

[72] Zanta-Boussif MA, Charrier S, Brice-Ouzet A, *et al.* Validation of a mutated PRE sequence allowing high and sustained transgene expression while abrogating WHV-X protein synthesis: application to the gene therapy of WAS. Gene Ther. 2009;16:605-619.

[73] Johnnidis JB, Harris MH, Wheeler RT, *et al.* Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. Nature. 2008;451:1125-1129.

[74] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126:663-676.

[75] Takahashi K, Tanabe K, Ohnuki M, *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131:861-872.

[76] Yu J, Vodyanik MA, Smuga-Otto K, *et al.* Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318:1917-1920.

[77] Kim JB, Zaehres H, Arauzo-Bravo MJ, Scholer HR. Generation of induced pluripotent stem cells from neural stem cells. Nat Protoc. 2009;4:1464-1470.

[78] Hanna J, Markoulaki S, Schorderet P, *et al.* Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. Cell. 2008;133:250-264.

[79] Rowland BD, Bernards R, Peeper DS. The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. Nat Cell Biol. 2005;7:1074-1082.

[80] Meng X, Neises A, Su RJ, *et al.* Efficient reprogramming of human cord blood CD34+ cells into induced pluripotent stem cells with OCT4 and SOX2 alone. Mol Ther. 2012;20:408-416.

[81] Liu T, Zou G, Gao Y, *et al.* High Efficiency of Reprogramming CD34(+) Cells Derived from Human Amniotic Fluid into Induced Pluripotent Stem Cells with Oct4. Stem Cells Dev. 2012.

[82] Kim D, Kim CH, Moon JI, *et al.* Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell. 2009;4:472-476.

[83] Papapetrou EP, Lee G, Malani N, *et al.* Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat Biotechnol. 2010;29:73-78.

[84] Sommer CA, Sommer AG, Longmire TA, *et al.* Excision of reprogramming transgenes improves the differentiation potential of iPS cells generated with a single excisable vector. Stem Cells. 2010;28:64-74.

[85] Mali P, Chou BK, Yen J, *et al.* Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. Stem Cells. 2010;28:713-720.

[86] Woltjen K, Hamalainen R, Kibschull M, Mileikovsky M, Nagy A. Transgene-free production of pluripotent stem cells using piggyBac transposons. Methods Mol Biol. 2011;767:87-103.

[87] Warren L, Manos PD, Ahfeldt T, *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell. 2010;7:618-630.

[88] Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc Jpn Acad Ser B Phys Biol Sci. 2009;85:348-362.

[89] Park IH, Arora N, Huo H, *et al.* Disease-specific induced pluripotent stem cells. Cell. 2008;134:877-886.

[90] Rashid ST, Corbineau S, Hannan N, *et al.* Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. J Clin Invest. 2010;120:3127-3136.

[91] Wang Y, Zheng CG, Jiang Y, *et al.* Genetic correction of beta-thalassemia patient-specific iPS cells and its use in improving hemoglobin production in irradiated SCID mice. Cell Res. 2012;22:637-648.

[92] Papapetrou EP, Lee G, Malani N, *et al.* Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat Biotechnol. 2011;29:73-78.

[93] Sebastiano V, Maeder ML, Angstman JF, *et al.* In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells. 2011;29:1717-1726.

[94] Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. Blood. 2011;118:4599-4608.

[95] Robbins J. Gene targeting. The precise manipulation of the mammalian genome. Circ Res. 1993;73:3-9.

[96] Delacote F, Lopez BS. Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model. Cell Cycle. 2008;7:33-38.

[97] Song H, Chung SK, Xu Y. Modeling disease in human ESCs using an efficient BAC-based homologous recombination system. Cell Stem Cell. 2010;6:80-89.

[98] Khan IF, Hirata RK, Wang PR, *et al.* Engineering of human pluripotent stem cells by AAV-mediated gene targeting. Mol Ther. 2010;18:1192-1199.

[99] Lombardo A, Genovese P, Beausejour CM, *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 2007;25:1298-1306.

[100] Gabriel R, Lombardo A, Arens A, *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol. 2011;29:816-823.

[101] Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods. 2011;8:765-770.

[102] Hockemeyer D, Soldner F, Beard C, *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009;27:851-857.

[103] Hockemeyer D, Wang H, Kiani S, *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011;29:731-734.

[104] Zavada J. VSV pseudotype particles with the coat of avian myeloblastosis virus. Nat New Biol. 1972;240:122-124.

[105] Cronin J, Zhang XY, Reiser J. Altering the tropism of lentiviral vectors through pseudotyping. Curr Gene Ther. 2005;5:387-398.

[106] Waehler R, Russell SJ, Curiel DT. Engineering targeted viral vectors for gene therapy. Nat Rev Genet. 2007;8:573-587.

[107] Peled A, Petit I, Kollet O, *et al.* Dependence of human stem cell engraftment and repopulation of NOD/SCID mice on CXCR4. Science. 1999;283:845-848.

[108] Yang L, Bailey L, Baltimore D, Wang P. Targeting lentiviral vectors to specific cell types in vivo. Proc Natl Acad Sci U S A. 2006;103:11479-11484.

[109] Joo KI, Wang P. Visualization of targeted transduction by engineered lentiviral vectors. Gene Ther. 2008;15:1384-1396.

[110] Lei Y, Joo KI, Wang P. Engineering fusogenic molecules to achieve targeted transduction of enveloped lentiviral vectors. J Biol Eng. 2009;3:8.

[111] Relander T, Johansson M, Olsson K, *et al.* Gene transfer to repopulating human CD34+ cells using amphotropic-, GALV-, or RD114-pseudotyped HIV-1-based vectors from stable producer cells. Mol Ther. 2005;11:452-459.

[112] Di Nunzio F, Piovani B, Cosset FL, Mavilio F, Stornaiuolo A. Transduction of human hematopoietic stem cells by lentiviral vectors pseudotyped with the RD114-TR chimeric envelope glycoprotein. Hum Gene Ther. 2007;18:811-820.

[113] Frecha C, Costa C, Negre D, *et al.* A novel lentiviral vector targets gene transfer into human hematopoietic stem cells in marrow from patients with bone marrow failure syndrome and in vivo in humanized mice. Blood. 2011;119:1139-1150.

[114] Lapillonne H, Kobari L, Mazurier C, *et al.* Red blood cell generation from human induced pluripotent stem cells: perspectives for transfusion medicine. Haematologica. 2010;95:1651-1659.

[115] Lu SJ, Feng Q, Park JS, *et al.* Biologic properties and enucleation of red blood cells from human embryonic stem cells. Blood. 2008;112:4475-4484.

[116] Ma F, Ebihara Y, Umeda K, *et al.* Generation of functional erythrocytes from human embryonic stem cell-derived definitive hematopoiesis. Proc Natl Acad Sci U S A. 2008;105:13087-13092.

[117] Hatzistavrou T, Micallef SJ, Ng ES, Vadolas J, Stanley EG, Elefanty AG. ErythRED, a hESC line enabling identification of erythroid cells. Nat Methods. 2009;6:659-662.

[118] Dias J, Gumenyuk M, Kang H, *et al.* Generation of red blood cells from human induced pluripotent stem cells. Stem Cells Dev. 2011;20:1639-1647.